Author:
Dewis, Lydia I

Title:
The Synergy of Synthesis, Computation and NMR Spectroscopy to Design Conformationally Controlled -Helix Mimetics

University of BRISTOL

The Synergy of Synthesis, Computation and NMR Spectroscopy to Design Conformationally Controlled α-Helix Mimetics

Lydia Dewis

Supervisors: Professor Craig P. Butts, Professor Varinder K. Aggarwal FRS, Professor Adrian J. Mulholland and Dr. Elisabetta Chiarparin

A dissertation submitted to the University of Bristol in accordance with the requirements for award of the degree of Doctor of Philosophy in the Faculty of Science.

University of Bristol – School of Chemistry

March 2020

Word Count: 77554

Abstract

In the search for high affinity therapeutics, medicinal chemists are turning their attention to free ligand conformations as it is well known that if a large conformational reorganisation is required upon binding, then a weaker binding affinity will be observed. This has been highlighted as an important strategy in the inhibition of protein–protein interactions (PPIs), due to the large, flat and featureless binding sites they present. This thesis discusses the design, synthesis and conformational analysis of a novel class of conformationally controlled α-helix mimetics, aimed at targeting aberrant PPIs. Their ability to target the p53-Mdm2 PPI with reasonable affinity is demonstrated.

Molecular Mechanics (MM) conformational search calculations were used to design a conformationally controlled scaffold that mimicked the *i*, *i* + 3/4 and *i* + 7 positions of the α-helix. Following the design of a general scaffold, appropriate side chains were added to mimic those of p53, and protein-ligand molecular docking calculations were performed to assess the binding of the designed ligand to Mdm2. Molecular docking was also used to design a library of p53 mimetics with binding affinities comparable to that of Nutlin-2, a highly successful p53-Mdm2 inhibitor.

With promising p53 mimetics designed using computation, their synthesis was explored using lithiation–borylation. Using lithiation–borylation a diverse set of side chains can be incorporated onto the scaffold by the insertion of the correct benzoate ester into the synthetic sequence. Thus, broad libraries of α-helix mimetics bearing different side chains can be synthesised using the same iterative methodology.

To confirm that the designed p53 mimetics exhibit the conformational bias that was predicted by MM calculations, a hybrid quantum mechanics (QM) and NMR spectroscopy approach was used to explore the conformation of one of the p53 mimetics. Experimentally measured scalar coupling constants, interproton distances derived from 1D-NOESY spectroscopy and both ^{1}H and ^{13}C chemical shifts were compared to the Boltzmann averaged values, calculated using DFT calculations. An excellent correlation was observed between the two data sets, confirming the predicted conformational bias.

Finally, the binding of the p53 mimetics to Mdm2 was explored using ^{1}H-^{15}N HSQC spectroscopy. The chemical shift perturbations (CSPs) that were observed upon binding of the ligands were mapped to amino acids located primarily in, or on the periphery of, the p53 binding pocket on Mdm2. The CSPs were tracked with increasing concentrations of ligand to obtain dissociation constants (K_d). Two different p53 mimetics were tested and binding affinities as low as 8.8 μM were obtained. Additionally, the binding of two control molecules was assessed using ^{1}H-^{15}N HSQC spectroscopy, to confirm the importance of both the hot-spot groups and conformational control in obtaining high affinity therapeutics.

Author's Declaration

"I declare that the work in this dissertation was carried out in accordance with the Regulations of the University of Bristol. The work is original, except where indicated by special reference in the text, and no part of the dissertation has been submitted for any other academic award. Any views expressed in the dissertation are those of the author.

SIGNED: DATE:

Acknowledgements

During my time at the University of Bristol, I have been fortunate enough to be supervised jointly by two exceptional supervisors, Professor Craig Butts and Professor Varinder Aggarwal FRS. I am eternally grateful to them both for their outstanding guidance and mentorship during this multidisciplinary PhD. Their continued support and encouragement enabled me to learn the diverse set of skills needed during my PhD, shaping me into the scientist I am today. I would also like to thank Craig for introducing me to the NMR community, which is without doubt the most collaborative and supportive group of scientists that I have met and immediately after I was immersed in this community I knew that NMR was the place for me. I would also like to thank my other supervisors Professor Adrian Mulholland and Dr. Elisabetta Chiarparin. Their helpful discussions and valuable suggestions helped to guide this project towards a successful conclusion. I am additionally grateful to Elisabetta for helping me secure my new position and I am excited to continue working with her at AstraZeneca.

During my PhD I had the opportunity to undertake a research internship at Genentech in San Francisco, and I would like to thank my supervisor Dr. Dave Russell and the structure elucidation group for making me feel so welcome. I had an incredible time and despite spending a considerable amount of time eating gelato with Dave, achieved some great results and published my first primary author paper! I would also like thank Dr. Jasmine Bickel and Laura Russell, fellow interns in a foreign land, whose friendship made the experience a truly unforgettable one.

Undertaking this multidisciplinary PhD has required collaboration with numerous scientists and I have had the opportunity to learn from a huge number of talented scientists across the University of Bristol. Firstly, I would like to thank Dr. Chris Williams whose guidance during the protein expression, purification and NMR was invaluable for the success of this PhD. I would also like to thank all of the members of the Aggarwal and Butts groups that I have had the pleasure of working with, particularly Dr. Beatrice Collins, Dr. Teerawut Bootwicha, Dr. Johan Pradeilles, Dr. Eddie Myers, Dr. Joe Bateman, Rory Cameron Mykura, Dr. Madhavachary Rudrakshula, Dr. Siying Zhong, Dr. Claire Dickson and least but not least the most tenacious organiser of pub trips and computer super nerd William Gerrard.

Finally, I would like to thank my friends and family who were always present during the highs and lows that one faces during a PhD. To Louise, James, Dabs and Jon, thank you for the many laughs we have shared during our PhDs. Kilian, mein Liebling, vielen Dank für alles. Du bist geduldig und verständnisvoll, und du findest immer einen Weg mich zum Lächeln zu bringen. To my sister Alexandra, thank you for teaching me to believe in myself, to challenge myself and to go after what I want. Finally, to my parents Jill and Gary, thank you for providing me with such exceptional opportunities. I cannot begin to summarise what your support and guidance has enabled me to achieve however, without it none of this would have been possible and so my biggest and most sincere thanks go to you.

Abbreviations and Acronyms

BMRB	Biological Magnetic Resonance Bank
CD	Circular Dichroism
COSY	Correlation Spectroscopy
CSA	Chemical Shift Anisotropy
CSP	Chemical Shift Perturbation
CSSF	Chemical Shift Selective Filter
DEPT	Distortionless Enhancement by Polarisation Transfer
DFT	Density Functional Theory
DIAD	Diisopropyl Azodicarboxylate
DQFCOSY	Double Quantum Filtered Correlation Spectroscopy
DPFGSE	Double Pulsed Field Gradient Selective Echo
ESI	Electrospray Ionisation
EXSIDE	Excitation Sculpted Indirect Detection Experiment
FA	Fluorescence Anisotropy
FP	Fluorescence Polarisation
GCMS	Gas Chromatography Mass Spectrometry
GIAO	Gauge Independent Atomic Orbitals
HCV	Hepatitis C Virus
HMBC	Heteronuclear Multiple Bond Correlation
HPLC	High Pressure Liquid Chromatography
HRMS	High Resolution Mass Spectrometry
HSQC	Heteronuclear Single Quantum Coherence
HSQMBC	Heteronuclear Single Quantum Multiple Bond Correlation
IEFPCM	Integral Equation Formalism Polarizable Continuum Model
IMAC	Immobilised Metal Affinity Column
INEPT	Insensitive Nuclei Enhancement by Polarization Transfer
IPAP	Inphase Antiphase
FTIR	Fourier Transform Infrared
ITC	Isothermal Titration Calorimetry
IPTG	Isopropyl β-d-1-thiogalactopyranoside
LB	Lysogeny Broth
MAD	Mean Absolute Deviation
MALDI	Matrix Assisted Laser Desorption Ionization
MCMM	Monte Carlo Multiple Minimum
MD	Molecular Dynamics
MM	Molecular Mechanics
MMFFs	Merck Molecular Force Field
MOM	Methoxy Methyl
MS	Mass Spectrometry
MST	Magnetic Shielding Tensor

NMR	Nuclear Magnetic Spectroscopy
NOESY	Nuclear Overhauser Effect Spectroscopy
OCb	Carbamate ester
OD	Optical Density
OPLS	Optimized Potential for Liquid Simulations
PANIC	Peak Amplitude Normalization for Improved Cross Relaxation
PDB	Protein Data Bank
PEG	Polyethylene Glycol
PPI	Protein–Protein Interaction
PSYCHE	Pure Shift Yielded by Chirp Excitation
QM	Quantum Mechanics
RCM	Ring Closing Metathesis
RMSD	Root Mean Square Deviation
ROESY	Rotating-Frame Nuclear Overhauser Effect Correlation Spectroscopy
SEC	Size Exclusion Chromatography
SBDD	Structure Based Drug Design
SD	Standard Deviation
SDS PAGE	Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
SOFAST HMQC	Band-Selective Optimised Flip Angle Transient Heteronuclear Multiple Quantum Coherence
STD	Saturation Transfer Difference
TBAI	Tetrabutylammonium Iodide
TCEP	Tris(2-carboxyethyl)phosphine
TIB	Triisopropylbenzoate
TLC	Thin Layer Chromatography
TMEDA	Tetramethylethylenediamine
TMS	Trimethyl Silane
TNCG	Truncated Newton Conjugate Gradient
TROSY	Transverse Relaxation Optimized Spectroscopy
TS	Transition State

Contents

Chapter 1: Introduction

1.1 Understanding and Controlling Molecular Conformation

The structure of a molecule, both its configuration and its conformation, are central to how the molecule behaves and the properties its exhibits. A different 2-dimensional or 3-dimensional arrangement of the same set of atoms will result in a unique molecule with distinct properties and thus correctly identifying a molecules 2D- and 3D- structure is crucial in molecular design.[1] Understanding how atoms are connected to form the 2D-structure is now relatively straightforward, owing to improvements in a number of analytical techniques, although none are as important as the developments in nuclear magnetic resonance (NMR) spectroscopy.[2] Examining a molecules 3D-structure or conformation, i.e. how the atoms arrange themselves in space once they are connected, is a much bigger challenge. The classic methodology used to confirm a molecules conformation or 3D-structure is X-ray crystallography. However, this technique requires a solid crystal of the molecule to be grown, which is not always a possibility. Additionally, only information regarding the solid-state conformation is provided, which can be affected by factors including crystal packing. Exploring the conformation of a molecule in solution is a much more challenging problem as conformational dynamics now come into play. Fortunately, NMR spectroscopy exists and can provide a plethora of information regarding molecular conformation in solution that one would want to ask, especially when coupled with computation.[3]

Section 1.1 will discuss the importance of molecular conformation, with a focus on the design of potent therapeutics, and discuss some of the methods that are available when trying to control molecular conformation. A short discussion of relevant examples from the literature demonstrating how molecular preorganisation can improve binding affinity is also present. The chapter will later go on to discuss how conformationally controlled molecules can be synthesised and analysed, before discussing the importance of conformational control in the inhibition of protein–protein interactions (PPIs).

1.1.1 The Importance of Molecular Conformation in Drug Discovery

In the drug discovery industry, structure-based drug design (SBDD) has fast become a mainstay in modern drug design and has led to the design of multiple successful small molecule drugs.[4–6] During this process, a small molecule is designed to fit complementary to its targeted binding pocket. Typically, x-ray crystal structures or NMR-derived three-dimensional structures of the receptor are used to design a ligand that fits into the binding pocket. If a structure of the receptor is not experimentally available, it is possible to computationally derive a model of the receptor.[7] Despite the success that has been achieved with SBDD, the process often overlooks two important points. The first being the conformation of the free ligand in solution and how accessible the bound conformation of the ligand

is.[8] Often the bound conformation of the ligand is not a low energy conformation in solution and thus the molecule must undergo a conformational rearrangement to fit into the binding pocket.[9] The second is that it assumes the receptor is rigid and that the ligand will dock by an 'induced fit' mechanism. However, this is not true, and a receptor can present many different conformations of the binding pocket to the ligand, one of which will be selected by the ligand that maximises binding affinity.[10] This PhD project is concerned with the former of these two issues and will explore methods to understand and control molecular conformation in solution.

Very often in SBDD the conformation of the free ligand in solution is not given adequate consideration during the drug design process. This is likely due to the lack of high throughput experimentation available that can accurately estimate the conformation of a flexible molecule in solution. Currently, the most accurate experimental method for analysing molecular conformation in solution is advanced NMR spectroscopy, such as NOESY spectroscopy, which is not particularly amenable to high-throughput experimentation.[11] This is unfortunate as it is known that if the conformation of the ligand free in solution does not reflect that of the bound ligand conformation a conformational reorganisation of the ligand is required for binding (**Figure 1**).[12]

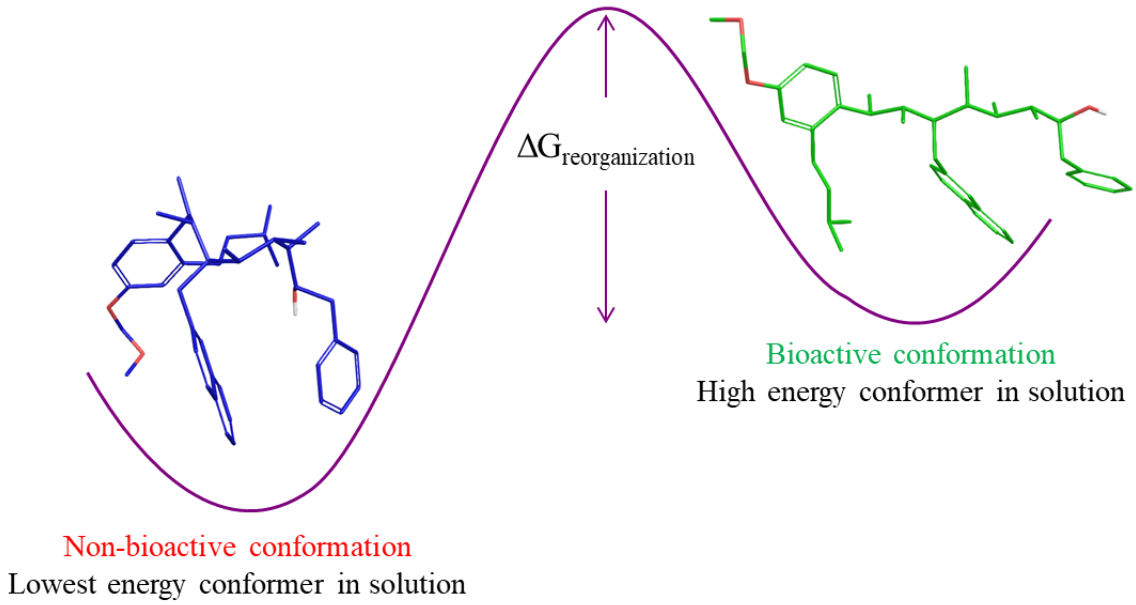


Figure 1. The conformational reorganisation required if the bound ligand conformation is not a low energy conformation in solution.

This conformational reorganisation reduces the overall binding affinity of the ligand to the receptor, due to both enthalpic and entropic factors.[6] Although it is experimentally not possible to measure the value of $\Delta G_{reorganization}$[13], a number of theoretical studies have been performed and values typically exist between 17-25 kJ mol^{-1}.[14] However, values as high as 109 kJ mol^{-1} have been reported.[15] Thus, flexible

ligands that can adopt multiple conformations, of which the bound conformation is not present as a low energy conformer, are not ideal drug candidates and will demonstrate lower binding affinities.

A more attractive approach is to design drug candidates rationally and intelligently, that display a conformational preorganisation of the free ligand in solution towards the bound conformation.[16–18] This prevents a conformational reorganisation from being required, lowering the entropic penalties of binding and thus increasing the binding affinity. A conformationally preorganised molecule also displays much faster on-rate kinetics (k_{on}) for protein-ligand binding, which also increases the binding affinity.[19] Since a conformationally preorganised molecule in solution effectively presents an increased concentration of the bioactive conformation towards the protein and no reorganisation is required prior to binding, an increase in the rate of k_{on} is observed.

It is clear that conformational preorganisation improves binding affinity and an understanding of free ligand conformation in solution can be used to rationally design ligands with higher binding affinities. If this analysis is completed during SBDD, it will significantly reduce the number of candidates having to be synthesised and tested, as a greater likelihood of finding a high affinity ligand will be achieved.[6] So, the question remains why this isn't standard practice in SBDD? As mentioned earlier it is very challenging to obtain experimental information regarding the free ligand conformation in solution, and typically advanced NMR spectroscopy and computational chemistry must be employed. Despite this, a number of examples in the literature exist from both industry and academia where an understanding of free ligand conformation in solution has been used to design conformationally preorganised molecules. A discussion of a small number of examples will be given in section 1.1.3, following a discussion of how molecular conformation can be controlled.

1.1.2 How is Molecular Conformation Controlled?

Before we can discuss how to control molecular conformation, an understanding of what conformational isomerism is will be helpful. Conformational isomers of a molecule arise from the free rotation around single bonds in the molecule to generate a new 3D- arrangement of the atoms, known as a conformer. The energies of the different conformers possible for a molecule differ based on various stereo- and electronic- factors.[20] The low energy conformers, or local minima on the energy surface, are referred to as conformational isomers of the molecule. The high energy conformations however, or local maxima, are referred to as transition states. The interconversion between conformations is dependent on the rotational energy barrier.[21] Controlling molecular conformation is typically concerned with making this rotational energy barrier higher for certain bond rotations, favouring certain conformations over others.

This concept of local minima and maxima existing on an energy surface can be demonstrated simply by looking at butane. The central C-C bond of butane can freely rotate giving rise to six different conformations with different energies, primarily differing due to steric interactions.[22] There are three local minima of butane (**1a**, **1c**, **1e**), all of which avoid sterically incumbering eclipsing interactions. These conformations are known as the staggered conformations. The global minimum, known as the *anti*periplanar conformation (**1a**), has a dihedral angle of 180° between the methyl groups, placing the methyl groups *anti* to one another. The two other staggered conformations (**1c** and **1e**) place the methyl groups in a *gauche* conformation, with a dihedral angle of 60°. The steric interactions between methyl groups in a *gauche* interaction increases the energy of the conformer by ~3.3 kJ mol^{-1}, calculated using *ab initio* Density Functional Theory (DFT) methods.[23] The three local maxima of butane (**1b**, **1d**, **1f**), all exhibit an eclipsed conformation. The global maximum (**1d**) features an eclipsed interaction between the methyl groups, increasing the energy of the conformer by 17-25 kJ mol^{-1}, due to both torsional strain and Me-Me steric interactions. The two other local maxima (**1b** and **1f**) feature a Me-H eclipsing interaction, raising the energy from the global minima by ~14.2 kJ mol^{-1}.[24,25] This conformational behaviour can be represented by a potential energy surface diagram (**Figure 2**).

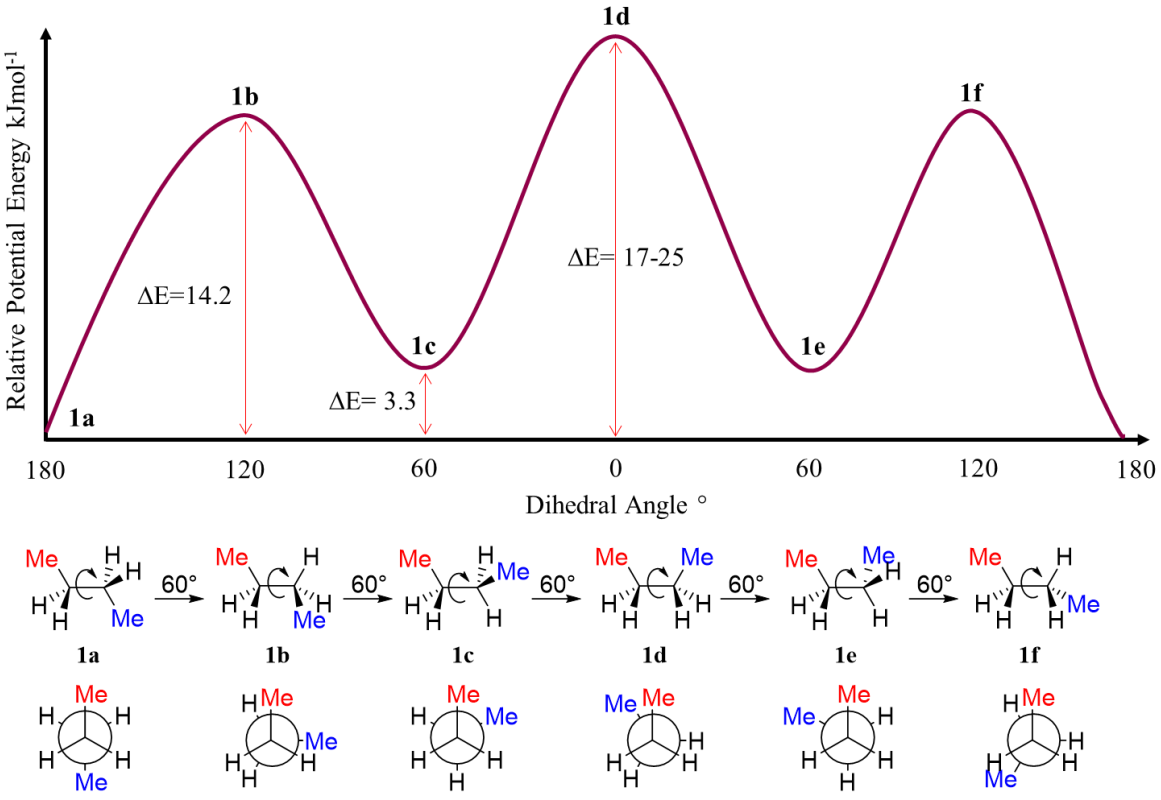


Figure 2. The potential energy diagram for butane. The Newman projections of the different conformations are shown below.

Unbranched alkyl chains feature complex conformational behaviour and have a huge number of low energy conformations populated at room temperature.[26] Although a global minimum will remain

present, typically the conformer will dihedral backbone dihedral angles of 180°, the propensity for the molecule to adopt this conformation reduces as the number of rotatable bonds increases because the enthalpic gain in reducing the number of unfavourable interactions is outweighed by the entropic cost of preorganisation.[27] However, it is possible to favour certain conformations over others by increasing the barrier to single bond rotation often through the introduction of steric strain. A discussion of some of the examples available for controlling conformation is included.

One method that has proven successful involves constraining acyclic linear chains into rings. For example, hexane has twelve unique conformations, which can freely interconvert at room temperature.[28,29] However, constraining hexane into a cyclic ring, and forming cyclohexane, reduces the number of rotatable bonds and as a result only six conformations of cyclohexane are possible, of which the two chair conformations are considerably lower in energy. A second example of introducing rings to control molecular conformation was described by Clark Still. He demonstrated that it was possible to control the conformation of the polyethylene glycol chain **2**, which has over 1000 conformations within 13 kJ mol^{-1} of the global minimum, by annealing six membered rings along the chain (**3).** This reduces the number of conformations that can exist within 13 kJ mol^{-1} of the global minimum to 25 (**Figure 3**).[1,30]

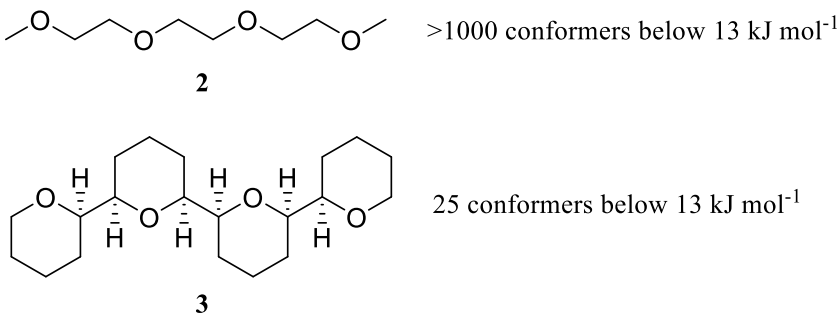


Figure 3. The effect of ring annellation on the conformational landscape of open chains.

The introduction of rings is a common strategy in the drug discovery industry, with many small molecule drugs containing multiple aromatic or aliphatic ring systems. It has also become increasingly common to constrain the conformation of flexible drugs with larger ring systems, through macrocyclization.[31,32] Although controlling molecular conformation through the introduction of rings has had considerable success, especially in the pharmaceutical industry, it is not without its limitations. Some of which include the addition of considerable molecular weight to the molecule and the fact that the introduction of rings is not always synthetically straightforward. A second strategy, that does not actively prevent the rotation of C-C bonds, but simply increases the barrier to rotation is the introduction of destabilising interactions.

A destabilising interaction that can control conformation or direct the outcome of stereoselective reactions is known as allylic strain, or $A^{1,3}$ strain.[33] The concept of allylic strain was first proposed by Johnson and Malhotra in 1965 when they noticed that certain conformations of cyclohexene were disfavoured over others due to unfavourable steric interactions for certain geometries of the double bond.[34] For example, consider the rotation around the single bond of an allylic system such as **4**. The global energy minimum for structures such as this, would be the one that places the two hydrogen atoms eclipsed to one another (**4a**), with a H-C-C=C dihedral angle between 0 ± 30°.[35] If two methyl groups are eclipsed however (**4b**), then $A^{1,3}$ strain is introduced forcing this conformation up in energy by ~17 kJ mol^{-1}, determined by ab initio DFT calculations (MP2-631G(d)//3-21G).[36] Although rotation of the vinylic bond is possible, it is unlikely at room temperature due to the high barrier of rotation enforced by $A^{1,3}$ strain. Thus, molecules such as **4** will typically populate a single conformation corresponding to **4a** (**Figure 4**).[33] The major factor governing the magnitude of allylic strain is the size of the substituents at the 1 and 3 positions, with large bulky substituents introducing greater allylic strain.[37]

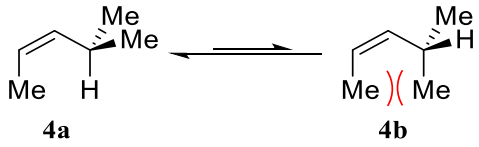
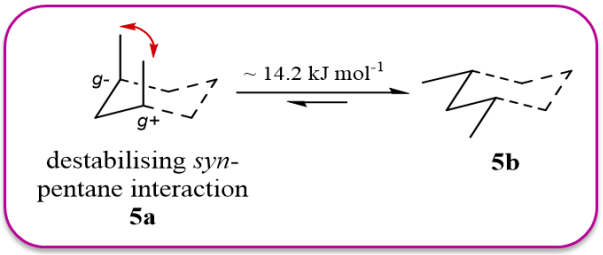
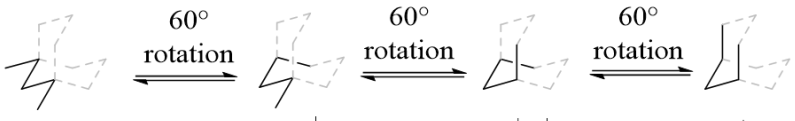


Figure 4. Conformational bias imposed by allylic strain.

A second destabilising interaction that can be introduced to a molecule to control conformation is the *syn*-pentane interaction.[38] Pentane, like butane, can adopt a number of conformations resulting from rotation of single bonds, and these conformations differ in energy (**Figure 5**). The lowest energy conformation of pentane is where the two dihedral angles are both *antiperiplanar,* eliminating any unfavourable interactions (*aa* **5b**). Rotation of one of the dihedral angles by 60° results in conformer **5c**, which contains an *antiperiplanar* dihedral angle followed by a *gauche*$^+$ dihedral angle (ag^+). The unfavourable *gauche* interaction increases the energy of conformer **5c** by ~3.3 kJ mol^{-1}, calculated using *ab initio* DFT calculations (MP4SDQ/6-31G(d)).[39] The rotation of the second dihedral angle by 60° gives a conformer featuring a *gauche*$^+$ dihedral angle followed by a second *gauche*$^+$ dihedral angle (g^+g^+ **5d**). Two unfavourable *gauche*$^+$ interactions are now present, increasing the energy of this conformation by ~5.9 kJ mol^{-1} from the global minimum. A final rotation of the first dihedral angle by 60° gives the global maximum, with a *gauche*$^+$ dihedral angle followed by a *gauche*$^-$ dihedral angle (g^+g^- **5a**). The two methyl groups are now eclipsed to one another and the hydrogen atoms on the carbon are in very close proximity resulting in a steric clash. This interaction is known as the *syn*-pentane interaction and it increases the energy of conformer **5a** by ~14.2 kJ mol^{-1} from the global minimum.[40]





	aa Anti-Anti **5b**	*ag*$^+$ Anti-Gauche **5c**	g^+g^+ Gauche-Gauche **5d**	g^+g^- *Syn*-Pentane **5a**
$\Delta\Delta H$ at 298.15 K (kJ mol^{-1})	0.0	3.3	5.9	14.2
Population at 298.15 K (%)	46.1	47.2	6.2	0.5

Figure 5. The different conformations of pentane and their relative energies

For pentane, the conformational control is poor with only one conformation being strongly disfavoured, however through the careful design of the substituent pattern on the hydrocarbon chain multiple *syn*-pentane interactions can be introduced to destabilise all but one conformer.[26] Controlling molecular conformation by exploiting the *syn*-pentane interaction was pioneered by Reinhard Hoffmann. He showed that the simple alkyl chain **6**, which would have numerous low energy conformations, can be controlled by the addition of methyl groups (**7** or **8**), inducing *syn*-pentane interactions (**Figure 6**).[33] The two molecules **7** and **8** have only two conformations that avoid a destabilising *syn*-pentane interaction. For molecule **7**, conformer **7a** is more energetically favourable, giving rise to a conformational bias towards an all-*trans* chain whereas for molecule **8**, conformer **8a** is more favourable leading to a bending of the main chain. This suggests that molecules of the type **7** and **8** have specific folding patterns due to the avoidance of *syn*-pentane interactions and that by changing the relationship of the methyl groups relative to each other it is possible to switch the shape of the backbone. This phenomenon has been observed previously in polymers of polypropylene.[41] Syndiotactic polypropylene, which features a repeating unit similar to **7** exists as an all-*trans* chain, or a linear chain. Isotactic polypropylene, which features a repeating unit similar to **8** exists as a helical chain with a change of direction observed at every other bond.

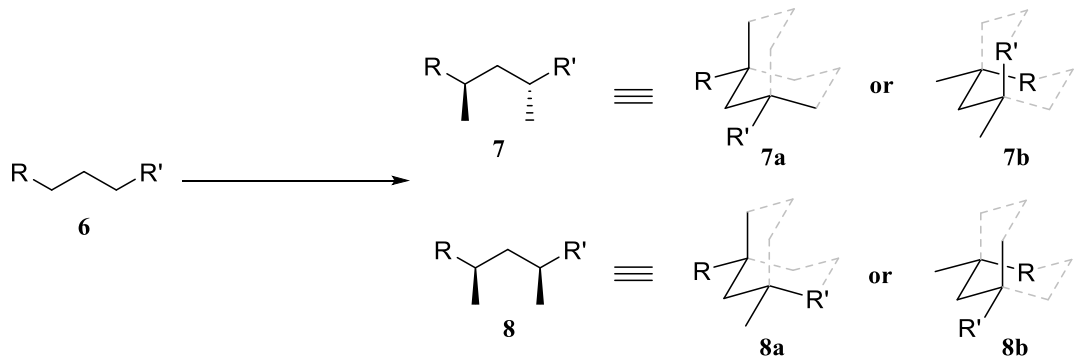


Figure 6. The conformational preference of **7** and **8**

Hoffmann also demonstrated that it is possible to reduce the number of conformations further from two to one by the introduction of another methyl group, to give polymethylated hydrocarbon chains. For example, on examination of 2,3,4,5-tetramethylhexane (**9**), only one conformation that is devoid of *syn*-pentane interactions is possible (**9a**) (**Figure 7**).[1] The rotation of any C-C bond in the molecule will create a *syn*-pentane rendering **9** monoconformational.

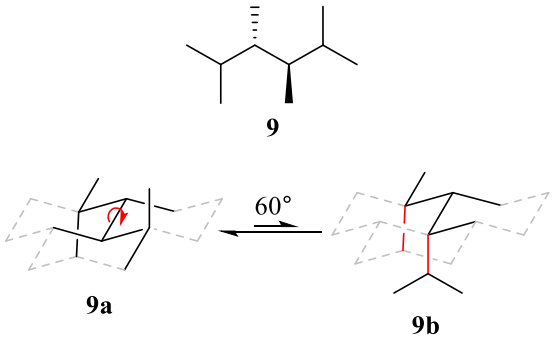


Figure 7. The conformational preference of 2,3,4,5-tetramethylhexane **9**

In 2014, Aggarwal *et. al.* demonstrated the utility of the *syn*-pentane interaction in controlling molecular conformation.[27] Inspired by the pioneering work of Hoffmann, Aggarwal *et. al.* synthesised molecules containing ten contiguous methyl groups on a hydrocarbon backbone. The molecules were synthesised by iterative lithiation–borylation[42], which will be discussed in section 1.2. Three different diastereomers were synthesised with differing relationships between the methyl groups. In an analogous fashion to molecules **7** and **8**, the molecules should take on defined conformations depending on the relationship of the methyl groups due to the avoidance of *syn*-pentane interactions.

The diastereomer that has all-*syn* relationships between the methyl groups (**10**) should adopt a helical conformation, while the diastereomer that has alternating *anti*-*syn* relationships between the methyl groups (**11**) should adopt a linear conformation (**Figure 8**). If any other conformation of these molecules were to be adopted, at least one *syn*-pentane interaction would be present, effectively reducing the population of these conformers to <0.5% at room temperature. The third diastereomer, the all-*anti*

diastereomer was predicted to have no conformational bias as no conformation of this molecule can avoid *syn*-pentane interactions. X-ray crystallography and NMR spectroscopy coupled with computation confirmed the suspected conformational bias for the three diastereomers, work which will be discussed in section 1.3.3. This work demonstrates how effective the incorporation of multiple *syn*-pentane interactions into a molecule can be at controlling molecular conformation.

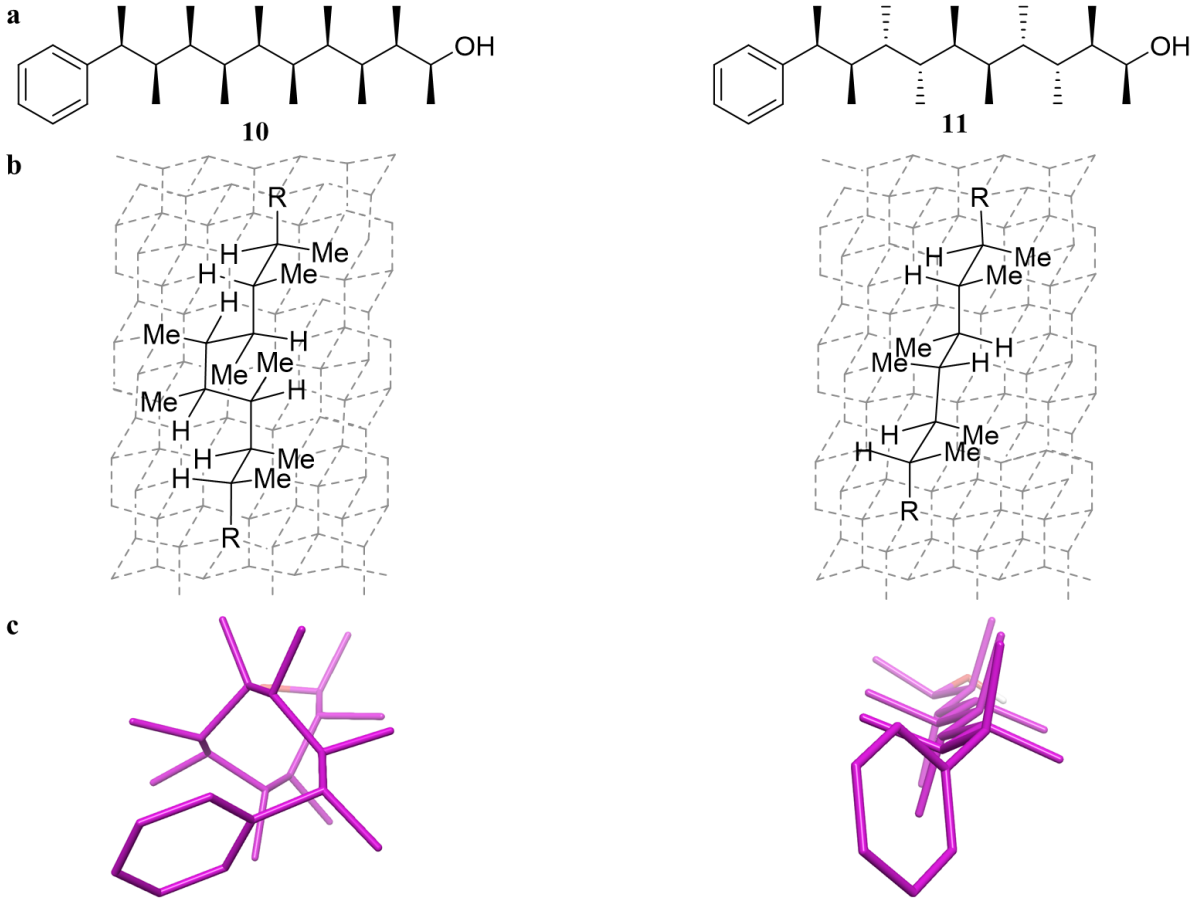


Figure 8. The conformational bias of **10** and **11**. (**a**) The structures of diastereomers **10** and **11**. (**b**) The diamond lattice analysis of the two diastereomers. Using the diamond lattice, it is clear that the rotation of any C-C bond would result in a conformer that has at least one *syn*-pentane interaction. (**c**) A 3D representation of **10** and **11**

1.1.3 Specific Examples of Molecular Preorganisation in Drug Discovery

Although controlling molecular conformation is not straightforward, and the analysis of free ligand conformations in solution is challenging, several examples exist in the drug discovery process where preorganised molecules which demonstrate improved binding affinities. A few of these examples will be discussed. For example, in 2017 a team of scientists from AstraZeneca reported the rational design of a high affinity Mcl-1 specific inhibitor.[43] Mcl-1 is a member of the Bcl-2 family of proteins and is involved in a number of protein–protein interactions (PPIs) that regulate apoptosis.[44] An initial affinity-mediated screen of DNA encoded libraries led to the discovery of **12** which bound to Mcl-1 with an

IC_{50} of 2 µM. Crystallisation of a derivative of **12** bound to Mcl-1 revealed that the bound conformation placed two sterically hindering groups in very close proximity, suggesting that the bound conformation is unlikely to be a low energy conformation of the free ligand in solution. They proposed that the free ligand conformation could be preorganised to adopt the bound conformation through macrocyclization. This resulted in compound **13** which demonstrated a much higher affinity to Mcl-1 with an IC_{50} of <3 nM. Indeed, the bound conformation of **13** matches with the free ligand conformation of **13**, predicted by NMR (**Figure 9**).[43]

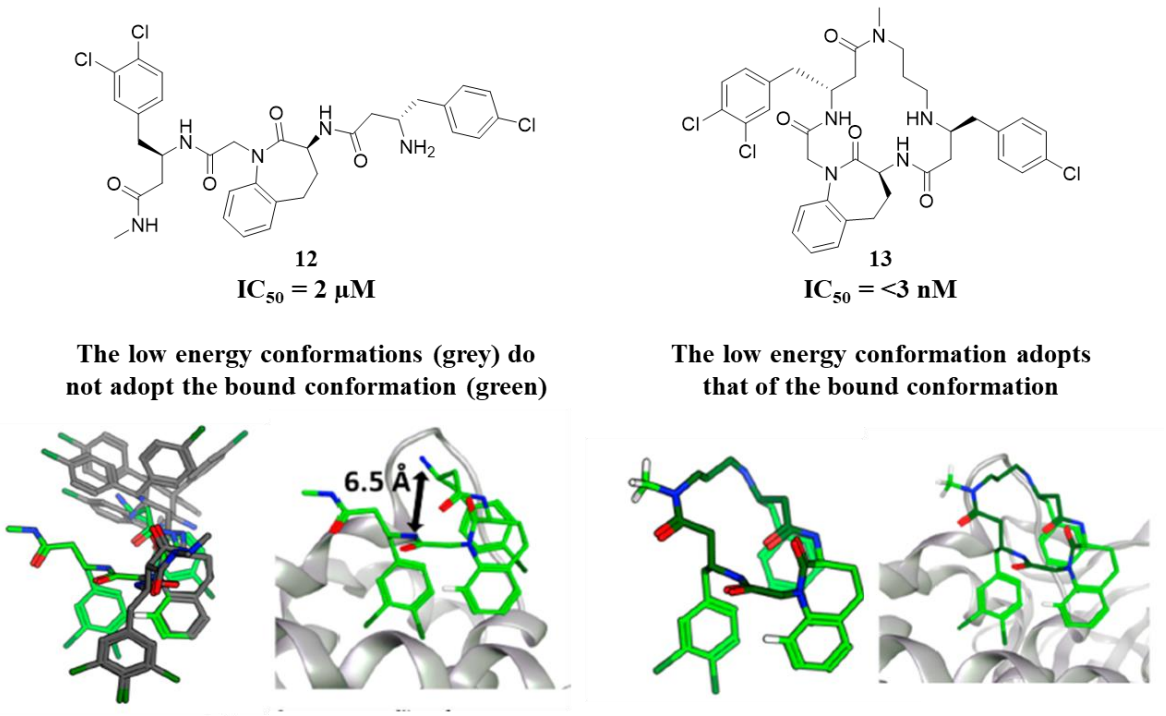


Figure 9. The 2D-structure and free and bound conformations of both **12** and **13**. The free ligand conformations of **12** are shown in grey, with the bioactive conformation superimposed in green (left). The free ligand conformation of **13** matches with the bound conformation of **13** (right). The figure was reproduced from reference[45] PDB ID: 5KU9 and 5MEV

Typically, a full analysis of scalar coupling constants, chemical shifts and interproton distances derived from NOESY is required to assess the conformational landscape of a flexible molecule in solution. However, the team of scientists at AstraZeneca demonstrated that the 1D ^{1}H-NMR spectrum could be used to indicate if the free ligand conformation adopted the bound conformation.[19] For the candidates discussed previously (**12** vs **13**), if the bioactive conformation was adopted then a large upfield shift of one of the aromatic residues is observed, caused by shielding from the dichlorosubstituted aromatic ring. The proton shown in blue on **13** (**Figure 10**), described as the reporter proton, shifts from 7.1 to 6.4 ppm after macrocyclization suggesting that the proton has been shielded by ring current effects from the dichlorosubstituted aromatic ring.[43] This would result in a conformation resembling that of the bioactive conformation. Additionally, upon rigidification from macrocyclization, each methylene pair (X, Y and Z) gives a unique chemical shift, indicative of conformational control.

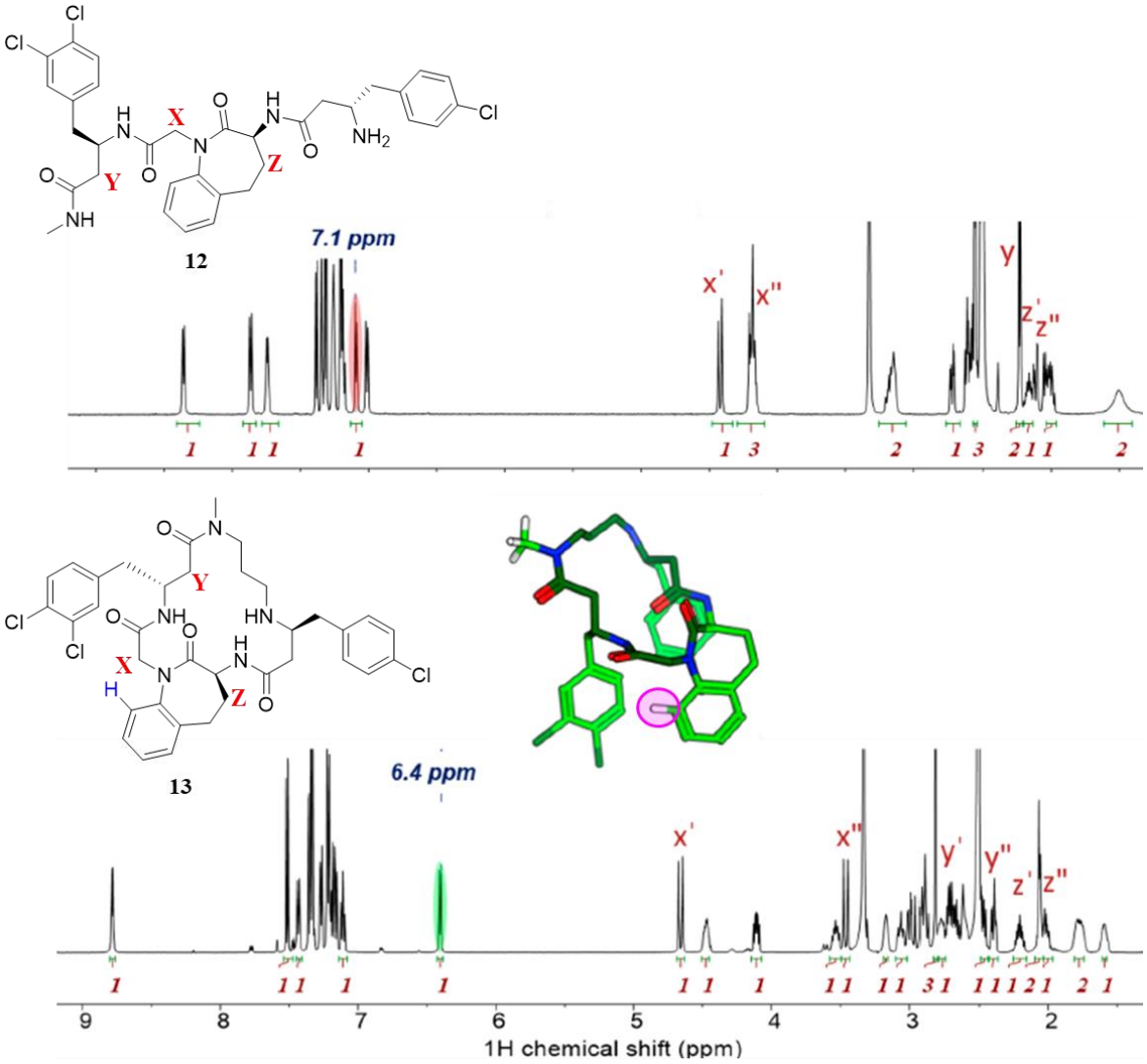


Figure 10. The 1D-^{1}H NMR spectrum of **12** (top) and **13** (bottom). The reporter proton is highlighted and the upfield shift from 7.1 ppm to 6.4 ppm is shown. The three methylene pairs indicative of macrocyclization, X, Y and Z, are also noted. The figure was reproduced from reference[45]

This methodology reported by AstraZeneca is incredibly powerful as the acquisition of 1D ^{1}H NMR spectra is very fast and the measurement of ^{1}H chemical shifts is straightforward. Using the ^{1}H NMR conformational signatures to indicate if the desired conformation is adopted, will enable the rapid screening of potential drug candidates. This information can be used to guide medicinal chemistry decisions, hopefully reducing the number of candidates that have to be synthesised.

Using a reporter proton and observing changes in the 1D-^{1}H NMR spectrum upon macrocyclization, AstraZeneca have been able to identify a number of preorganised macrocyclic drugs, all of which show an improved binding affinity compared to their flexible counterparts.[10,46] Macrocyclization is a well-known strategy in ligand rigidification and as a result the prevalence of macrocycles in small molecule drug design is rising.[47,48] However, macrocyclization is not limited to the conformational control of

small molecules. It has also found use in the conformational control of peptides. Peptide therapeutics, despite their typically higher selectivity and specificity over small molecules, often demonstrate poor conformational control and proteolytic digestion.[49,50] Maintaining the secondary structure of short peptide sequences is challenging and typically they display significant conformational flexibility in water. This limits their biological activity and thus strategies that control the conformation of peptides are of high importance.

One strategy that has proven successful in the control of peptide conformation involves the introduction of a covalent staple across between 4 to 7 residues of the peptide sequence, forming a macrocycle across the peptide.[51] Typically, olefinic side chains are added to judicious residues in the peptide sequence and the macrocycle is formed by ring closing metathesis (RCM).[52,53] Such peptides are known as stapled peptides and they are very efficient at stabilising the secondary structure of a peptide, essentially locking it in a certain conformation, whilst also helping to improve the pharmacokinetic properties by introduction of certain groups onto the macrocycle (**Figure 11**).[54] A huge number of stapled peptides have been reported in the literature, almost all of which show improved binding affinity to their desired target that the unstapled version.[51]

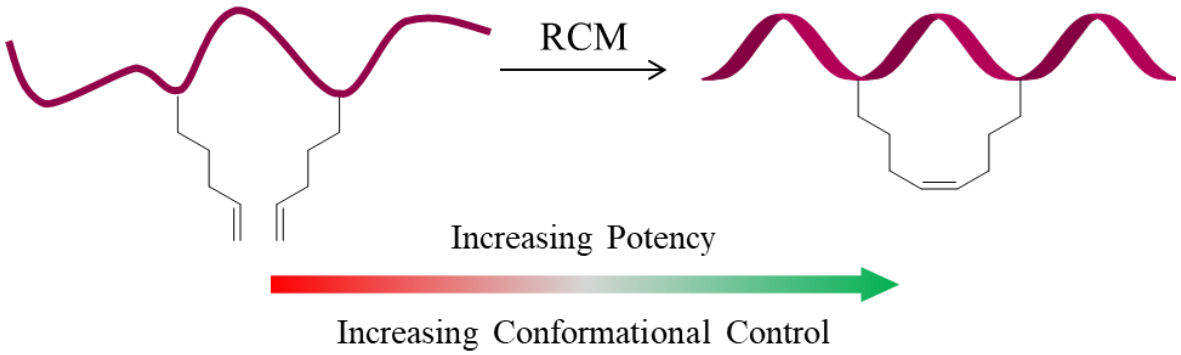


Figure 11. A schematic to show the impact on peptide structure upon introducing a staple.

In 2016, Keating *et. al.* demonstrated that by increasing the helicity of Mcl-1 inhibitor peptides through both stapling and the introduction of residues known to promote helicity results in an increased binding affinity attributed to slower off-rate kinetics.[55] The introduction of a hydrocarbon staple (M1d) across the original peptide sequence (MS1) increases the % helicity by 3-fold, determined by CD spectroscopy. The % helicity was increased further by introducing non-natural amino acid residues with helix inducing side chains (M2d and M3d). A positive linear correlation between % helicity and binding affinity was observed (**Figure 12a** and **Figure 12b**). Additionally, a positive linear correlation between % helicity and peptide half-life was also observed, suggesting that the peptides with a higher % helicity were also more stable to protease digestion (**Figure 12a** and **Figure 12b**). This work clearly demonstrates the importance of conformation in obtaining higher affinity therapeutics inhibitors.

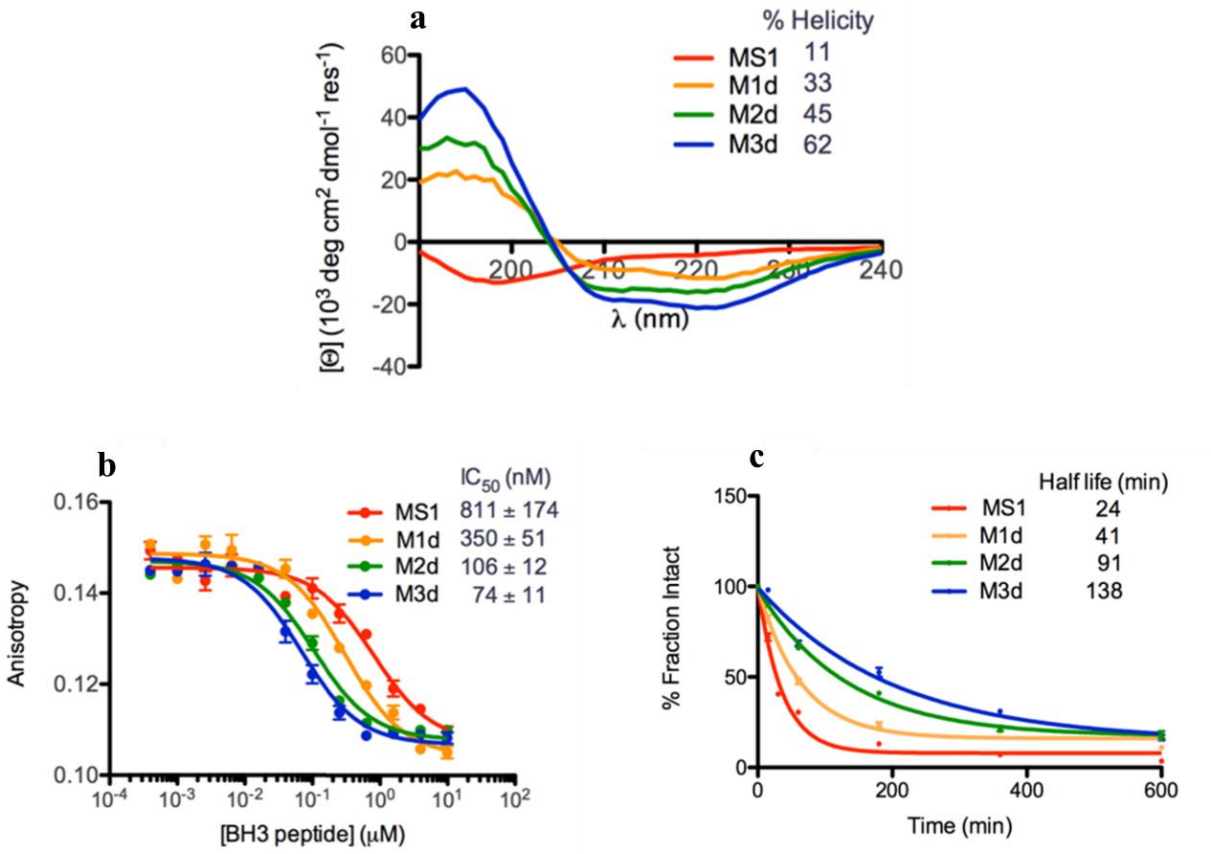


Figure 12. The results obtained from Keating *et. al.* clearly demonstrate that a high % of helicity is required for potent binding. **a)** The results from the CD spectroscopy which show an increase in helicity from the unstapled peptide MS1 to the stapled peptides M1d – M2d. **b)** The IC_{50} obtained for the four different peptides showing an increase in potency as the helicity is increased. **c)** The increase in half-life as the helicity is increased in clearly demonstrated on exposure of the peptides to chymotrypsin. This figure was reproduced from reference[55]

In the previous example, the peptide sequences were very long, with >20 amino acids in the sequence. However, much shorter peptide sequences are also common therapeutics, with a molecular weight and behaviour more closely aligned to a traditional small molecule. The hepatitis C virus (HCV) protease enzyme has been a long sought-after target of the pharmaceutical industry and nearly all strategies in the literature are derived from a hexapeptide sequence (DDIVPC, **14**).[56] However, since it is well known that peptidic therapeutics display complex conformational behaviour, a group of scientists at Boehringer Ingelheim, led by Stephen Kawai, used the structure of DDIVPC to design a conformationally controlled derivative.[57]

The team of scientists at Boehringer Ingelheim first confirmed by NMR spectroscopy that the free conformation of DDIVPC did not reflect that of the bound conformation. Following this, a number of modifications were made to the structure of DDIVPC, employing classical SBDD methods, including truncation of the N-terminus to yield a tri-peptide moiety, introduction of a *t*-butyl group and introduction of a cyclopentene ring, to yield the inhibitor **15** (**Figure 13**). This inhibitor showed a

significantly improved binding affinity compared to DDIVPC. A crystal structure of a urea analogue of **15** bound to the active site of the HCV protease enzyme was obtained (PDB: 4K8B), confirming that **15** bound in a very similar fashion as DDIVPC. ROESY NMR spectroscopy was used to obtained quantitative interproton distances of **15** in solution, which suggested that the free conformation of **15** in solution correlated very well with that of the bound.

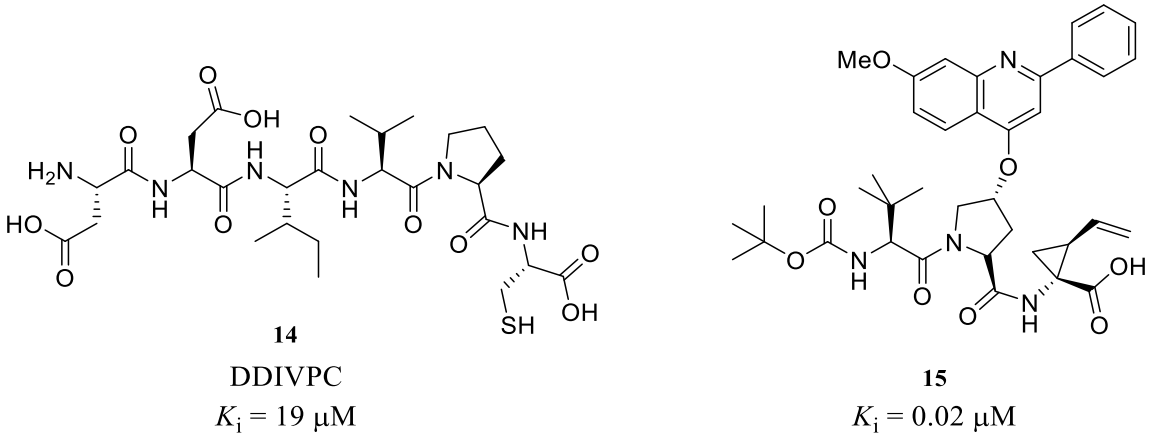
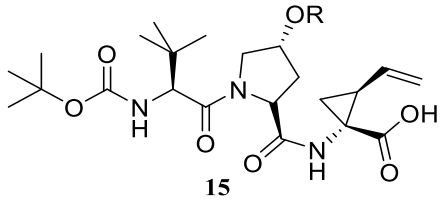
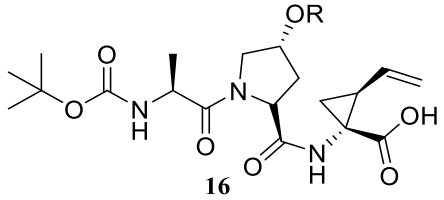
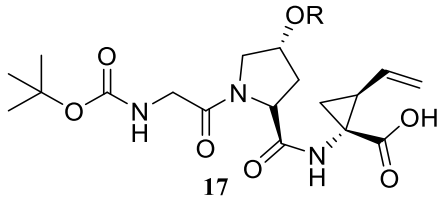


Figure 13. The structure and binding affinities of DDIVPC (**14**) and **15**.

The improvement in the binding affinity observed from DDIVPC to **15** supports the theory that by conformationally restraining the free conformation towards the bound conformation it is possible to increase binding affinity, However, the group at Boehringer Ingelheim used NMR spectroscopy, specifically $^3J_{NH\text{-}H\alpha}$ coupling constants, to confirm that this is indeed true for this example. They synthesised three derivatives of **15**, with decreasing steric bulk on the side chain of first amino acid. They found that as they decreased the steric bulk from *t*-butyl>methyl>proton, the value of $^3J_{NH\text{-}H\alpha}$ also decreased from 8.3>7.6>5.6 Hz (**Table 1**). For structure **15**, the *t*-butyl group enforces a *trans* conformation of the peptide bond, resulting in a large coupling constant. For structure **17**, the hydrogen is not capable at controlling conformation and allows free rotation around the peptide bond, resulting in a smaller coupling constant. The positive effect of the conformational control exerted by the *t*-butyl group, and also the methyl group to a lesser extent, is observed as an improvement in the binding affinity. Thus, this work serves to further confirm the importance of exploiting conformational control in the search for high affinity therapeutics.

Table 1. The effect on the binding affinity upon reducing the steric bulk on the side chain of first amino acid. The decrease in binding affinity is caused by a reduction in the number of conformers adopting the required *trans*-conformation, evident from the decrease in the $^{3}J_{NH\text{-}H\alpha}$ coupling constant

Inhibitor	K_i (μM)	$^{3}J_{NH\text{-}H\alpha}$ (Hz)	% *trans*-conformer
 **15**	0.02	8.3	87
 **16**	0.31	7.6	80
 **17**	3.0	5.6	41

The previous two examples discussed the conformational control of peptide therapeutics, however, peptides suffer significantly from proteolytic digestion and poor membrane permeation and are thus often considered undesirable. In an attempt to overcome this, peptoids were developed.[58] The term peptoid was coined to describe synthetic oligomers based on repeating units of oligo(*N*-substituted glycine) (oligo-NSG, **19**).[59–61] Peptoids typically display enhanced proteolytic stability and membrane permeability than peptides, however they still suffer from poor conformational control. Attempts at controlling the conformation of peptoids was previously limited to the introduction of bulky and hydrophoboic N-substituents[62,63] however, this severely limits the water solubility and restricts the choice of the N-substituent. Kodadek *et. al.*[64] and Sando *et. al.*[65] have addressed this issue through the introduction of a methyl group on the backbone α-carbon of oligo-NSG, synthesising what is known as oligo(*N*-substituted alanine) (oligo-NSA, **20**) (**Figure 14**).

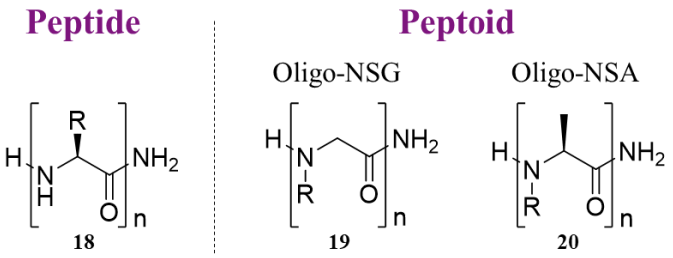


Figure 14. The structures of a peptide, an oligo-NSA and an oligo-NSG

The methyl group introduced causes conformational control primarily through the avoidance of pseudo-1,3-allylic strain. Two mechanisms by which pseudo-1,3-allylic strain is avoided are operational. The first being between the methyl group and the carbonyl oxygen of the preceding residue and the second between the methyl group and the N-substituent of the next residue. These two mechanisms restrict the rotation around the dihedral angles φ and ψ respectively. Additionally, the rotation around a third dihedral angle, ω, is restricted to ~180° to avoid a steric clash between the methyl groups that would exist if the value of ω was ~0°. This forces dihedral angle ω, which exists between the peptide bond, into a *trans* conformation (**Figure 15**). These mechanisms result in oligo-NSA having an extended linear conformation in water. It was also found that the nature of the N-substituent did not affect the extended conformation of oligo-NSA and the N-substituents were well-dispersed in space.[65]

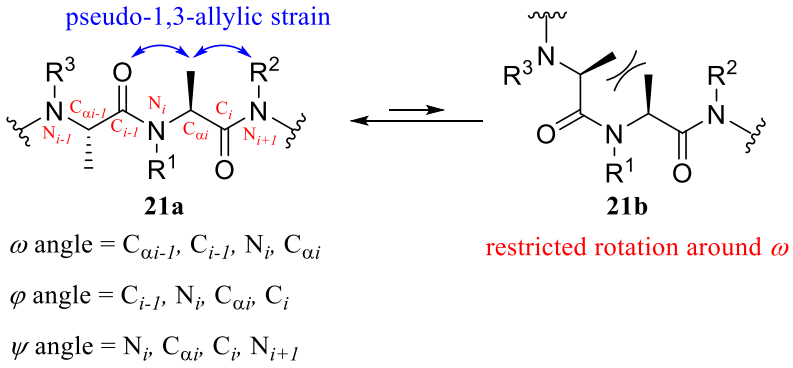


Figure 15. A diagram showing the definition of the three important dihedral angles ω, φ and ψ and the pseudo-1,3-allylic strain and restricted rotation around ω controlling the conformation of oligo-NSA.

Sando and co-workers demonstrated that oligo-NSA has superior biological activity to oligo-NSG, due to its conformational control, through the design of an inhibitor of the p53-Mdm2 PPI. The p53-Mdm2 PPI is perhaps the most well-known PPI in the scientific literature and is involved with cellular pathways regulating apoptosis to DNA repair.[66] More about this PPI will be discussed in section 1.5. Sando *et. al.* synthesised a p53 mimetic based on both oligo-NSG and oligo NSA and reported a binding affinity to Mdm2 of >300 μM and 1.1 μM respectively.[65] This clearly demonstrates that by conformationally preorganising a molecule, even one as large as a peptoid, much higher binding affinities can be obtained.

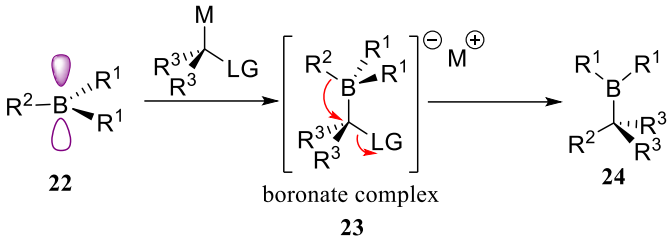
Many of the examples that have been discussed in this section have involved the design of an inhibitor of a PPI. PPIs are very challenging targets in drug discovery, for reasons which will be discussed in section 1.4, however conformational rigidity of potential small molecule inhibitors has been highlighted as a key parameter to consider in their design, highlighted recently in a review by Lawson *et. al.*[67] This PhD project aims to further demonstrate that conformational control of a small molecule PPI inhibitor can improve the binding affinity to these notoriously hard to drug targets.

1.2 The Iterative Homologation of Boronic Esters

Synthesising conformationally controlled molecules is not a straightforward process. Often the introduction of new stereocentres are required, and enantioselective synthesis is challenging. For example, controlling molecular conformation through the avoidance of destabilising *syn*-pentane interactions requires contiguously substituted methyl groups to be placed enantiospecifically onto a hydrocarbon backbone. The next section will discuss a synthetic methodology that enables the synthesis of contiguously methyl substituted hydrocarbons.

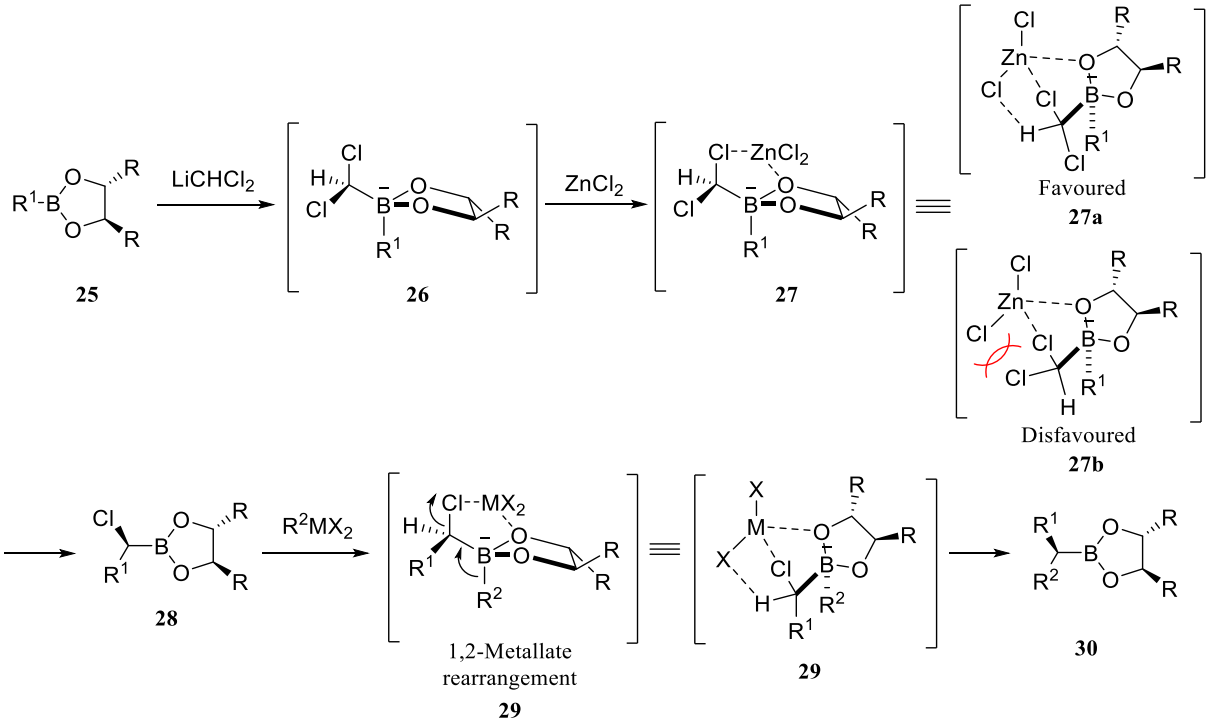
Organoboranes and boronic esters are one of the most synthetically versatile groups and can be transformed into a wide range of functional groups, often with complete stereospecificity.[68] Their impressive reactivity is due to the vacant boron-centred p-orbital which is able to accept a nucleophilic electron pair to form boronate complexes. Organoboranes are very electrophilic, owing to this vacant p-orbital, and can react with a wide range of nucleophiles however, they typically display very poor stability and are prone to oxidation. Boronic esters circumvent many of these stability issues due to π-donation from the lone pair of an oxygen atom of the ester group into the vacant p-orbital of boron.[69] This π-donation does reduce the electrophilic character of boronic esters, compared to organoboranes, however their improved stability renders them the preferred functional group in organoboron chemistry.

One of the most attractive features of organoboranes and boronic esters is that they can be homologated cleanly and with high stereochemical fidelity, provided there is a leaving group on the nucleophilic carbon and that a substituent on boron can arrange itself antiperiplanar to the leaving group.[42] If this is the case, the boronate complex **23** can undergo a stereospecific 1,2-metallate rearrangement to yield a homologated species **24**, with inversion of stereochemistry (**Scheme 1**). During this process, a boron bound group migrates to the adjacent carbon atom with the subsequent loss of the leaving group.[70]



Scheme 1. The reaction of an organoboron species with a nucleophile to form a boronate complex which can undergo a stereospecific 1,2-metallate rearrangement

The homologation of boronic esters was first reported in 1980 by Matteson in which a chiral boronic ester reacts with (dichloromethyl)lithium to give a homologated boronic ester (**Scheme 2**).[71,72] In this approach, a chiral auxiliary is attached to the diol group of the boronic ester and thus the chirality of the product is determined by the substrate. This approach is known as the substrate-controlled homologation of boronic esters.



Scheme 2. The substrate-controlled homo logation of boronic esters proposed by Matteson.

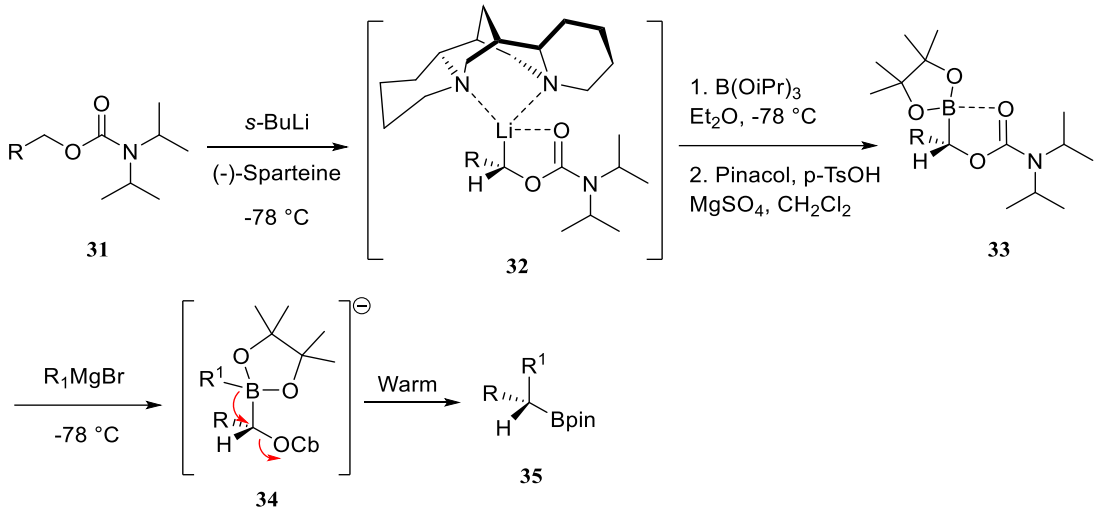
The first step of the substrate-controlled homologation of a boronic ester involves the reaction of an enantiopure boronic ester **25** with (dichloromethyl)lithium to yield a (dichloromethyl)borate complex **26**. The subsequent addition of zinc chloride causes the borate complex to undergo a rearrangement via the favoured transition state **27a** to give an α-chloroboronic ester (**28**) in high diastereomeric purity, typically >99:1 dr. The zinc chloride is thought to promote this rearrangement by interaction with the less hindered oxygen and assisting with the loss of a chloride ion. It is also proposed that the zinc

chloride stabilises the preferred TS (**27a**) through interaction of a chloride from the zinc chloride with the C-H bond of the borate complex. The newly formed α-chloroboronic ester **28** can react with a Grignard or organolithium reagent to give a borate complex **29**. The borate complex **29** then undergoes a stereospecific 1,2-metallate rearrangement to yield a homologated boronic ester, **30**.[73]

Although Matteson's substrate-controlled homologation methodology is able to homologate a boronic ester with exceptional stereocontrol, the stereochemistry of the product is governed by the stereochemistry of the chiral auxiliary attached to the boronic ester.[42] In order to access a different stereoisomer a sometimes challenging manipulation is required to invert the stereochemistry of the diol on the boronic ester. This can add up to three steps to the sequence, making this an unattractive approach for the iterative homologation of a boronic ester. A more attractive approach would be for the stereochemical outcome to be dictated by the stereochemistry of the reagent. This would enable access to different stereoisomers by simply switching the enantiomer of the reagent used. This approach is known as the reagent-controlled homologation of boronic esters.

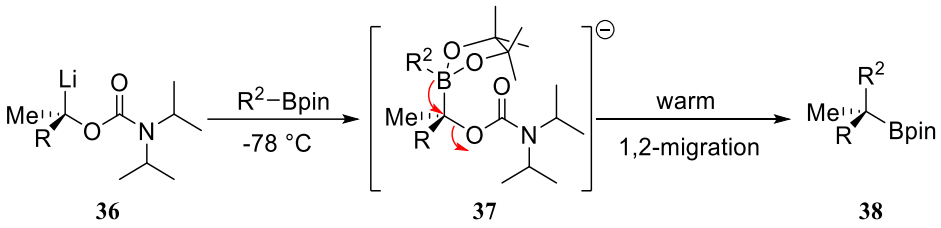
The reagent-controlled homologation of boronic esters involves the formation of a chiral boronate complex, using a suitable chiral reagent, that can then undergo a stereospecific 1,2-metallate rearrangement to yield a homologated chiral boronic ester that can then be used in subsequent homologations. In order to achieve this, a reagent that can act as a chiral carbanion that also possesses a leaving group at the same carbon as the carbanion, is required.[74] Additionally, the chiral carbenoid needs to display excellent chemical and configurational stability.[75]

One class of compounds that meets these requirements are Hoppe's lithiated carbamates. Hoppe has shown that chiral lithiated carbamates (**32**), generated from the chiral ligand-assisted deprotonation of a carbamate (**31**), can be trapped with $B(OiPr)_3$ to yield an α-carbamoyl alkylboronate, which can be converted to the corresponding boronic ester through the addition of pinacol borane (**33**) (**Scheme 3**).[76,77] Hoppe showed that both enantiomers of the chiral diamine sparteine could be used to generate both enantiomers of the chiral carbanion upon deprotonation with *s*-BuLi. The corresponding chiral carbenoid was chemically and configurationally stable at -78 °C. Addition of a Grignard reagent to the newly formed boronic ester gave a boronate complex (**34**) that, upon warming the flask to room temperature, underwent a 1,2-metallate rearrangement to yield a homologated boronic ester (**35**) after expulsion of the carbamate moiety as a leaving group. The carbamate group (OCb) also acts as a directing group, directing the deprotonation through the formation of a pre-lithiated species.[78]



Scheme 3. Hoppe's alkylation of boronic esters using lithiated carbamates

Following on from this seminal work by Hoppe *et. al.*, the groups of both Aggarwal[75] and Kocienski[79] have shown that a pinacol boronic ester (or borane) can be added directly to the lithiated carbamate (**Scheme 4**). The methodology developed by both groups has now been coined lithiation–borylation.

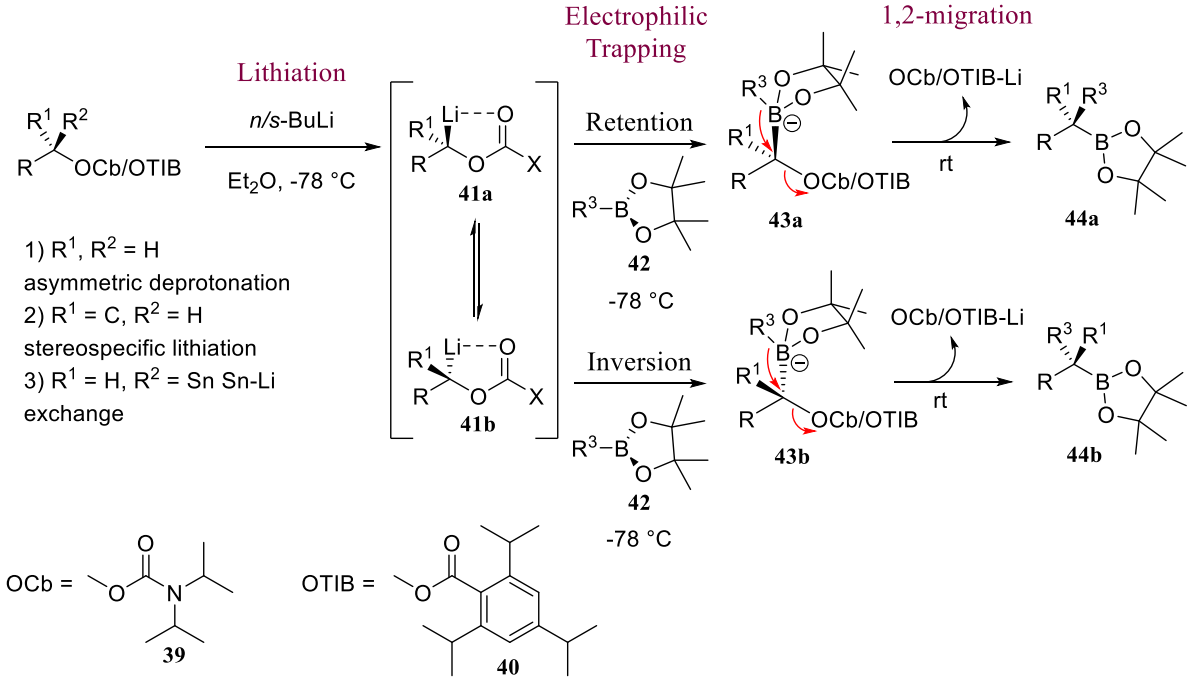


Scheme 4. The direct reaction of a pinacol boronic ester with enantioenriched lithiated carbamates developed by both Kocienski and Aggarwal.

Whilst exploring the scope of this reaction, Aggarwal *et. al.* found that certain boronic esters were slow to migrate and Lewis acids were required to promote the 1,2-metallate rearrangement. Because of this, alternative leaving groups that could enhance the rate of the 1,2-metallate rearrangement were considered. Beak has previously reported that 1,4,6-triisopropylbenzoates (TIB) can be deprotonated by *s*-BuLi and TMEDA at the α position, suggesting that TIB esters would be suitable in lithiation–borylation.[80,81] Aggarwal *et. al.* have reported that using the hindered TIB esters in lithiation–borylation enables the 1,2-metallate rearrangement of slow migrating groups.[82]

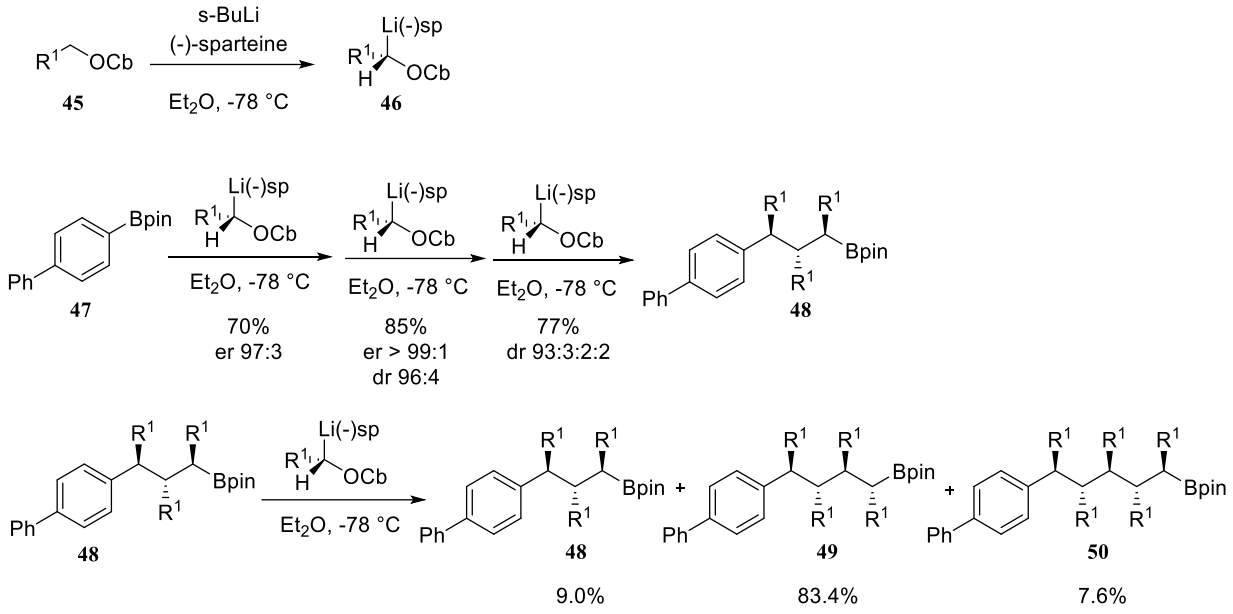
The mechanism of lithiation–borylation involves three distinct steps: lithiation, electrophilic trapping or borylation, and the 1,2-metallate rearrangement (**Scheme 5**). The first step involves the formation of a chemically and configurationally stable chiral carbenoid (**41a/b**) through the α-lithiation of either a carbamate (**39**) or a benzoate (**40**). The chiral carbenoids can be accessed using the chiral ligand-assisted asymmetric deprotonation of primary carbamates or benzoates. Alternatively, the stereospecific

deprotonation of enantioenriched secondary carbamates or benzoates can be used. Finally, the chiral carbenoids can be accessed from the stereospecific tin-lithium exchange of enantioenriched stannanes. Following this, an electrophilic organoboron species (**42**) traps the newly formed chiral carbenoid to form a chiral boronate species (**43a/b**). Again, this species is chemically and configurationally stable at -78 °C. Typically this process occurs with retention of configuration however, invertive pathways have been observed when using benzylic carbamates and boranes. The final stage is a stereospecific 1,2-metallate rearrangement of a group from boron to its attached carbon, resulting in expulsion of a leaving group. This forms a new carbon-carbon bond and yields the homologated boronic ester (**44a/b**). This process only occurs upon warming the flask to room temperature.



Scheme 5. The mechanism of lithiation–borylation

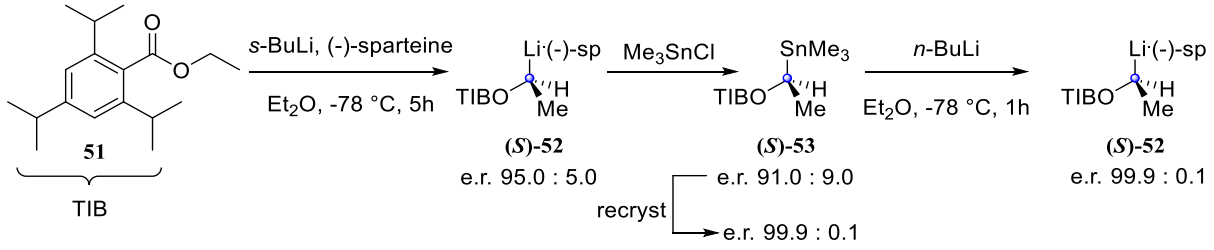
Aggarwal *et. al.* have gone on to show that this process can be performed iteratively, with the homologated boronic ester being used as the starting boronic ester in a subsequent homologation. Using Hoppe's lithiated carbamates, they demonstrated that three iterative homologations of a boronic ester could be performed in good yield and good *d.r.* Further homologations were less successful with a mixture of starting boronic ester (under homologated product), desired product and the over homologated product being formed (**Scheme 6**). They proposed that the 1,2-metallate rearrangement of the boronate complex was too slow, allowing it to reverse back to the starting boronic ester and the lithiated carbamate.[83] This would allow for **48** to react with any homologated product **49** that had formed to generate the over homologated product **50**.



Scheme 6. The three iterative homologations performed by the Aggarwal group using lithiation–borylation

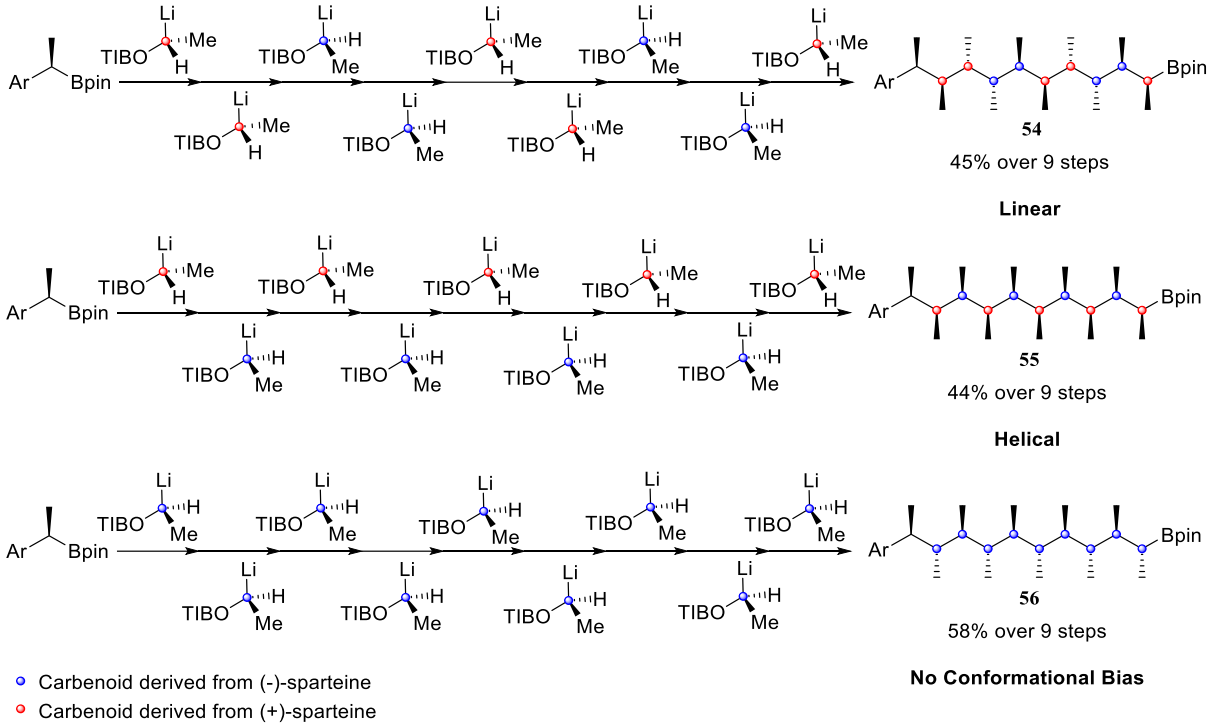
To overcome this problem, Aggarwal *et. al.* explored the use of triisopropylbenzoates in the iterative homologation of boronic esters, as they have been shown to aid the rearrangement of slow migrating groups. They demonstrated that, provided a chiral carbenoid could be accessed in >99.9:0.1 *e.r.*, the iterative homologation of boronic esters could be performed to generate molecules with up to ten contiguous methyl substituents with total stereocontrol.[27] If the chiral carbenoid cannot be accessed in high enantiomeric ratios and if each homologation does not occur with yields in excess of 95% then the resulting product distribution after nine iterative homologations would be extremely complicated.

In order to meet these requirements, Aggarwal *et. al.* used the stereospecific tin-lithium exchange of enantioenriched stannanes for generation of the chiral carbenoid (**Scheme 7**). The enantioenriched stannanes (**53**) can be generated from the lithiation of ethyl triisopropylbenzoate (**51**) in the presence of *s*-BuLi and (±)-sparteine followed by electrophilic trapping with trimethyltin chloride to yield the α-stannyl benzoate (**53**) in good enantiopurity. The resulting stannane can be recrystalised in methanol to provide the high enantiopurity required. Using this methodology, both enantiomers of the stannane were able to be synthesised as bench stable crystalline solids, ready for use in the iterative process.



Scheme 7. Generation of the α-lithiated hindered benzoates in high enantiomeric ratio from the corresponding stannanes.

Using the enantioenriched stannanes as the chiral carbenoid precursor, nine iterative homologations could be performed with essentially one diastereomer formed. This resulted in the preparation of hydrocarbons bearing ten contiguously substituted methyl groups with complete stereocontrol and no column chromatography required until the end. Three diastereomers were synthesised (**54**, **55**, **56**) by changing the enantiomer of the stannane used in the sequence (**Scheme 8**). This process has been likened to a molecular assembly line of molecules with tailored shapes. By synthesising molecules of this type, the Aggarwal group were able to confirm the *syn*-pentane control that is expected for these molecules, as discussed in section 1.1.2. The conformational bias that was expected of the three diastereomers was confirmed using a combined NMR spectroscopy and computational approach.[27] This strategy for examining molecular conformation will be discussed in the next section.



Scheme 8. The iterative assembly line synthesis of three diastereomers containing ten contiguously substituted methyl groups

1.3 The Analysis of Molecular Conformation using a hybrid NMR Spectroscopy and Molecular Mechanical/Quantum Mechanical Computational Approach

Investigating the conformational landscape of a molecule in solution is not a trivial challenge, especially for molecules with a large number of rotatable bonds such as molecules **54**, **55** and **56**. NMR spectroscopy is arguably the most useful analytical tool for exploring conformation in solution as it provides information about the conformational ensemble of species that are present in solution. For a rigid molecule, with only a few conformers contributing to the conformational landscape at room temperature, NMR spectroscopy alone can be used to estimate conformer populations. However, for a flexible molecule with complex conformational dynamics contributing to the NMR spectra observed, NMR spectroscopy alone cannot accurately estimate conformer populations. Coupling NMR spectroscopy with molecular mechanics (MM) and quantum mechanics (QM) calculations can overcome this problem and provide insights into the molecular conformation in solution for highly flexible molecules.

This hybrid NMR spectroscopy and computational approach has been routinely used throughout the literature to study the conformation of a molecule in solution and in the assignment of relative stereochemistry.[84–86] During this process, the experimentally derived chemical shifts, scalar coupling constants and interproton distances are compared to those calculated by density functional theory calculations (DFT). The mean absolute deviation (MAD) obtained between calculated and experimental data can be compared to the MAD typically observed in the literature to assess whether the computation has correctly estimated the conformational landscape of the molecule in solution.[3] To use this method to assign relative stereochemistry, it is possible to calculate the NMR parameters of the possible diastereomers and compare to the experimental NMR data. The correct diastereomer should provide the lowest MAD between experimental data and calculated data.

This section will discuss the hybrid NMR and QM workflow used in the analysis of solution conformation. Select examples from the literature that have employed the hybrid QM and NMR spectroscopy approach to solving 3D molecular structure or relative stereochemistry will also be discussed.

1.3.1 Extracting 2D- and 3D- Molecular Information using NMR Spectroscopy

NMR spectroscopy provides a plethora of information regarding chemical structure and conformation. It is well known that the ^{1}H and ^{13}C chemical shift provides information regarding the chemical environment and that multiplicities of peaks can deduce how many protons are adjacent to the one in

question. This can help to build up a 2D-structure of the molecule however, much more information can be extracted using NMR spectroscopy that can also help deduce a molecules 3D-structure.

The first parameter that will be discussed are vicinal scalar coupling constants, or ^{3}J coupling constants, which have a strong dependence on the dihedral angle between the coupling nuclei. This was first reported by Martin Karplus in 1963.[87,88] The relationship between ^{3}J and dihedral angle, described by the Karplus equation, and summarised in the Karplus curve is routinely used in the elucidation of 3D molecular structure and has proved reliable for a wide range of molecules. However, care must be taken if a molecule exhibits atypical bond lengths or bond angles, caused by ring strain for example, or contains electronegative substituents in the coupling pathway.[89–91] The equation was first developed to describe the relationship of 3-bond ^{1}H-^{1}H vicinal coupling constants ($^{3}J_{HH}$) to dihedral angle, however it has since been shown that a similar relationship is observed between 3-bond ^{1}H-^{13}C vicinal coupling constants ($^{3}J_{HC}$) and dihedral angle.[92]

The relationship between $^{3}J_{HH}$ and $^{3}J_{CH}$ with dihedral angle be clearly seen from the Karplus curve shown in **Figure 16**. Using the Karplus relationship, a dihedral angle of 180°, corresponding to an *anti*-relationship between X and H, will result in a large coupling constant ($^{3}J_{HH}$ = 8 – 14 Hz, $^{3}J_{CH}$ = 6 – 8 Hz), due to optimal orbital overlap between the vicinal sp^3 orbitals. However, a dihedral angle of 60°, corresponding to a *gauche* relationship between X and H, will result in a small coupling constant ($^{3}J_{HH}$ = 2 – 5 Hz, $^{3}J_{CH}$ = 1 – 3 Hz). Thus, the magnitude of the coupling constant can be used to deduce the dihedral angle between the vicinal nuclei, helping to elucidate a 3D structure of a molecule.[93]

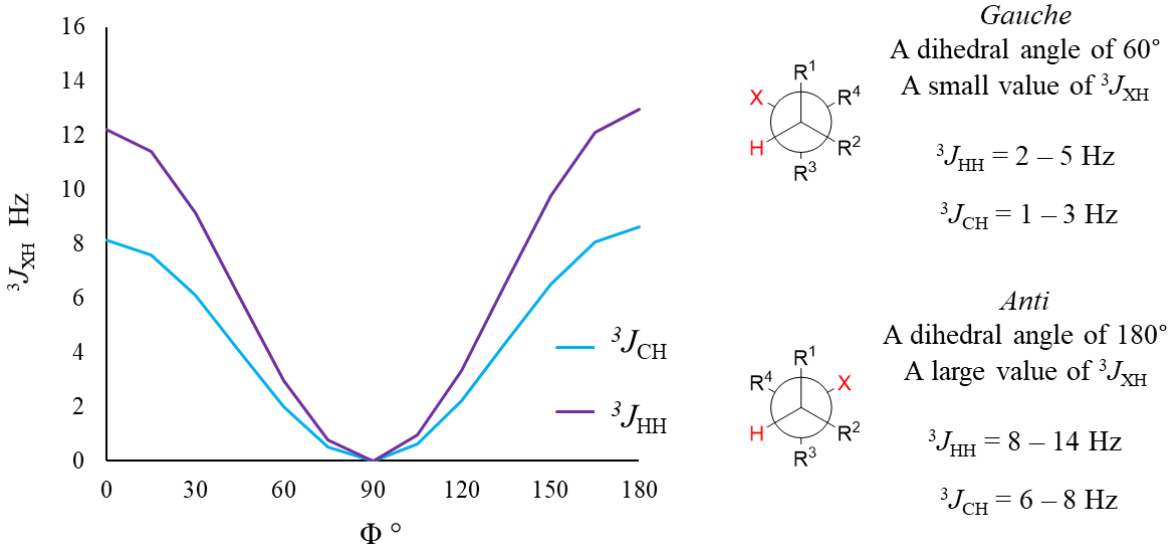


Figure 16. The relationship between dihedral angle and vicinal scalar coupling constants.

The experimental measurement of $^{3}J_{HH}$ is typically a straightforward process, with couplings measured directly from well-resolved signals in the 1D ^{1}H NMR spectrum. For overlapped regions of the spectra,

a number of 2D-methods exist which can aid in the extraction of $^{n}J_{HH}$ values. For example, 2D J-resolved methods, such as the 2D-J PSYCHE, which decouple spectra in the direct dimension (F_2) and projects the multiplets to a second dimension (F_1), from which the multiplet structure can be extracted.[94] However, the acquisition of 2D J-resolved spectra is not trivial and is often plagued by poor lineshapes resulting from a 45° titling of the final 2D spectra.[95] Thus, it is very advantageous to have well resolved signals in the 1D ^{1}H spectrum, from which the values of $^{3}J_{HH}$ can be measured.

It is also possible to simulate complex or overlapped multiplets using spin simulation software such as SpinSimulation from MestreNova. Using spin simulation, the simulated multiplet arising from estimated coupling constants is compared to the experimental coupling constant. If deviations between the two multiplets are observed then the simulated multiplet is re-simulated using altered coupling constants and the process is repeated until no deviations are observed between the two multiplets. This method prevents the acquisition of additional spectra, however for severely overlapped or broad peaks it is usually still insufficient for measuring $^{3}J_{HH}$ values.

Although $^{3}J_{HH}$ values provide valuable conformational information, they provide no information about quaternary centres within the molecule. This is where the measurement of $^{3}J_{CH}$ values can complement the measurement of $^{3}J_{HH}$ and provide further conformational information. In addition to providing information regarding quaternary centres, there are typically many more $^{3}J_{CH}$ values in a molecule than $^{3}J_{HH}$. Furthermore, $^{3}J_{CH}$ values can pass through heteroatoms and provide information connecting different spin systems. Despite the wealth information than $^{3}J_{CH}$ values can provide, they are not routinely used in conformational analysis due to the challenges associated with their measurement, primarily caused by the low natural abundance of ^{13}C.[96–98] Large values of $^{n}J_{CH}$ (typically, $^{1}J_{CH}$) can be directly measured from the 1D ^{1}H spectrum because they appear as ^{13}C satellites. However, the intensity of these peaks is very weak and often overlapped with other ^{1}H resonances in the spectrum. Values of $^{n}J_{CH}$ can also be obtained from 1D coupled ^{13}C spectra, however the spectra often comprise of complex overlapped multiplets caused by large values of $^{1}J_{CH}$. Additionally, the spectra take a considerable amount of time to acquire due to direct ^{13}C-detection.[99] Nevertheless, a number of 2D- methods exist that have enabled to accurate measurement of $^{n}J_{CH}$ values.[98] In 2017 Dickson *et. al.* reviewed the popular 2D-methods, focussing on the accuracy and practicality of the various methods, and reported that for a global $^{n}J_{CH}$ analysis of a molecule the IPAP-HSQMBC (Inphase-Antiphased hetereonuclear single quantum multiple bond correlation) was the best choice.[99] This is primarily due to the ease of extraction of coupling constants from the data and the reduced experimental time required to measure the $^{n}J_{CH}$ values (~3-6 minutes per $^{n}J_{CH}$, whilst other methods commonly require at least three times this).

In a IPAP-HSQMBC experiment two separate HSQMBC spectra are recorded, one with the $^{n}J_{CH}$ scalar couplings in-phase and one antiphase with respect to ^{13}C. The two spectra can either be recorded as

separate experiments or interleaved. The sum and the difference of the IP and AP FID's are performed prior to the Fourier transformation of each. The sum and difference spectra are then overlaid and the $^{n}J_{CH}$ scalar couplings can be extracted from the offset between multiplets in the F2 dimension between the two spectra (**Figure 17**).[100,101] The IPAP-HSQMBC experiment makes it simpler to extract couplings as it does not rely on the lineshape analysis of complex F2 multiplets. However, it is reliant on the sum and difference spectra exhibiting the same multiplet shape, which is not always the case. Additionally, the value of $^{n}J_{CH}$ measured from the different peaks of an overlaid multiplet is not always consistent. This can be addressed by calculating the $^{n}J_{CH}$ value as an average of the splitting's.[99]

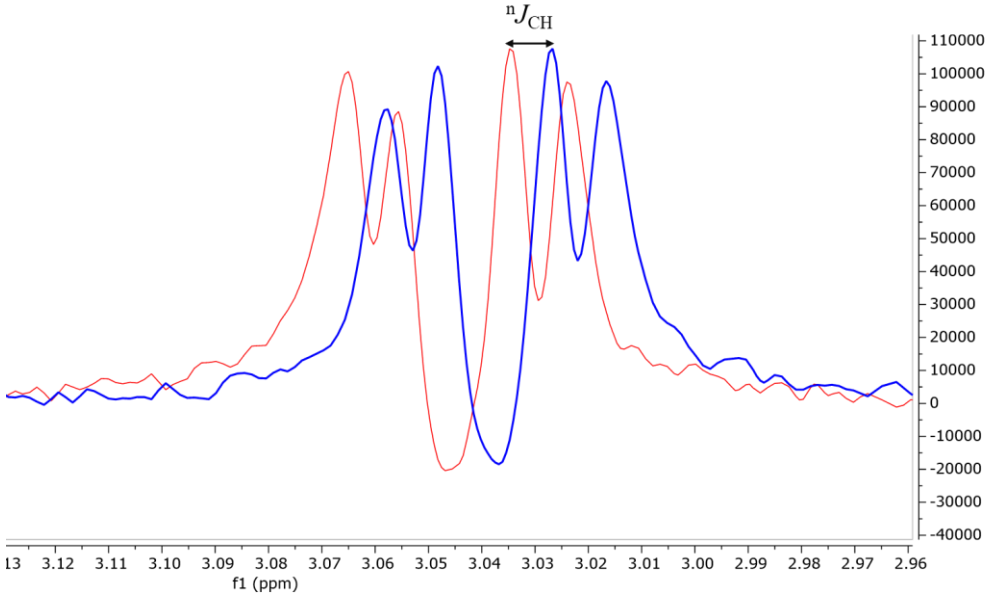


Figure 17. The IPAP-HSQMBC analysis of a multiplet. The sum (red) and the difference (blue) of the IP and AP HSQMBC spectra are overlaid and the coupling is measured from the offset between the multiplets.

Another experimental NMR technique that provides useful 3D-information of a molecule is NOESY (Nuclear Overhauser Effect SpectroscopY). In a NOESY experiment, spin active nuclei that are close in space (typically <5 Å) can transfer spin polarization between dipolar coupled nuclei, giving rise to an NOE correlation. The strength of the NOE correlation has a $1/r^6$ dependence on the internuclear distance and thus NOESY can be used to provide information on the spatial arrangement of nuclei.[102] Typically, the nuclei concerned are protons, however heteronuclear NOE between, for example ^{1}H and ^{19}F, has also been reported.[103] This methodology has historically been performed in a qualitative (NOE/no NOE) or semi-quantitative (weak/medium/strong) manner only, due to a lack of accuracy associated with a quantitative NOE analysis largely caused by additional cross-relaxation pathways contributing to the observed NOE.[104] However, improvements in NOE experimental methods and data analysis have demonstrated that quantitative NOE analysis is in fact accurate.

The NOE experiment can be used to obtain quantitative information regarding interproton distances due to the relationship between NOE intensity, cross relaxation rate and mixing time (**Equation 1**). The observed NOE intensity between two spins *I* and *S* (η_{IS}) at a given mixing time (τ_m), is proportional to the cross-relaxation rate between I and S (σ_{IS}).

$$\eta_{IS} = \sigma_{IS}\tau_m$$

Equation 1

The cross-relaxation rate between *I* and *S* is, in turn, inversely proportional to the internuclear distance between *I* and *S* (r_{IS}^{-6}) to the sixth power, when in the fast tumbling regime (**Equation 2**). In the fast tumbling regime, NOEs are independent of rotational correlation time (τ_c) and Larmor frequency and are dominated by the double quantum (ω_2) cross relaxation pathway and are thus positive NOE correlations. Therefore, assuming that the parameters which define *k* remain constant, the NOE intensity between spins *I* and *S* (η_{IS}) is proportional to the internuclear distance between the two nuclei (r_{IS}). This however, assumes that the initial rate approximation is true and that NOE build up is linear with mixing time. If this is the case, it is therefore true that the ratio of NOE intensities for a pair of NOE signals is proportional to the ratio of their internuclear distances. Thus, using a well-defined rigid reference distance (r_{ref}) that is not affected by conformational dynamics, the distance between other nuclei in the molecule (such as *I* and *S*) can be determined by comparing the relative NOE intensities of the peak of interest (η_{IS}) to that of the reference distance (η_{ref}) (**Equation 3**). A common reference distance is that between geminal protons of a methylene group (~1.78 Å).

$$\sigma_{IS} = kr_{IS}^{-6} \quad \text{Where } k = \left(\frac{\mu_0}{4\pi}\right)\frac{\hbar^2\gamma^4}{10}\left(\frac{6\tau_c}{1+4\omega^2\tau_c^2} - \tau_c\right)$$

Equation 2

$$r_{IS} = r_{ref}\left(\frac{\eta_{IS}}{\eta_{ref}}\right)^{-\frac{1}{6}}$$

Equation 3

However, NOE intensities do not build up linearly with mixing time, and deviations away from the initial rate approximation typically occur at short mixing times (~250 ms). A linear build up between NOE intensity and mixing time is usually only true at short mixing times when the NOE intensity is weak and the quantitative analysis of cross-relaxation rates is less accurate due to higher signal to noise.[105] The non-linearity seen at longer mixing times is largely attributed to external relaxation processes.

In 1986, Macura *et. al.* reported the quantitative analysis of 2D-NOESY spectra by internal calibration of the cross-peak NOE intensities against the diagonal peak intensities.[106] This approach effectively extends the length of mixing times that have a linear correlation with NOE intensity, enabling the measurement of cross-relaxation rates to be performed at longer mixing times which benefit from increased signal intensity. The effects of external relaxation are in effect cancelled out in this approach.

This approach was extended by Krishnamurthy to selective 1D-NOE experiments and the method was coined PANIC (Peak Amplitude Normalization for Improved Cross Relaxation).[107] The PANIC method corrects for different forms of relaxation by standardization of the irradiated peak which allows for NOE intensities obtained in separate inversion experiments to be directly compared. This effectively eliminates any perturbations to the irradiated peak or cross-peaks that proportionately affect all spins in an NOE experiment. The PANIC method successfully eliminates any nonlinear NOE buildup, which is typically observed at moderate mixing times, and thus the ratio of NOE intensities for a pair of NOE signals is proportional to the ratio of their internuclear distances and importantly this ratio is independent of mixing time. Therefore, by using the PANIC method, longer mixing times can be used which increases the signal-to noise and significantly improves the accuracy of the analysis.[108]

Butts *et. al.* have since demonstrated that this method can be used to extract experimental interproton distances from both rigid[104] and flexible[109] molecules, with deviations as low as 3% when compared to either QM calculated structures or x-ray crystal structures. They have also demonstrated that this method can be used to identify conformers with populations as low as 2%.[102] NOE is much more sensitive to high energy conformations, and traditional NMR methods to explore conformation such as scalar couplings would not have been able to identify conformers with populations this low.

For example, in 2011, Butts *et. al.* used quantitative 1D-NOESY to identify a high energy conformation of strychnine (**57**), previously unidentified using standard conformational analysis. Their initial quantitative analysis of strychnine using NOESY gave a MAD of 3.5% and a standard deviation (SD) of 2.7% when the experimental NOE distances were compared to those from DFT calculation and X-ray crystallography.[104] Although the overall correlation was good, an unusually large deviation of 14.9% was observed for the H_{11b}-H_{23b} distance. The experimentally derived interproton distance (3.49 Å) was shorter than the distance observed in the DFT calculations (4.10 Å), suggesting that a high energy conformer with a shorter H_{11b}-H_{23b} exists. If this were true, this conformation would have a disproportionally large effect of the NOE intensity because of the strong r_{IS}^{-6} relationship between distance and NOE intensity. Further computational calculations revealed a second conformer with a population of 2.5% at room temperature with a H_{11b}-H_{23b} distance of 2.11 Å (**Figure 18**). Accounting for this conformer in their analysis, reduced the MAD to 2.8%.[102]

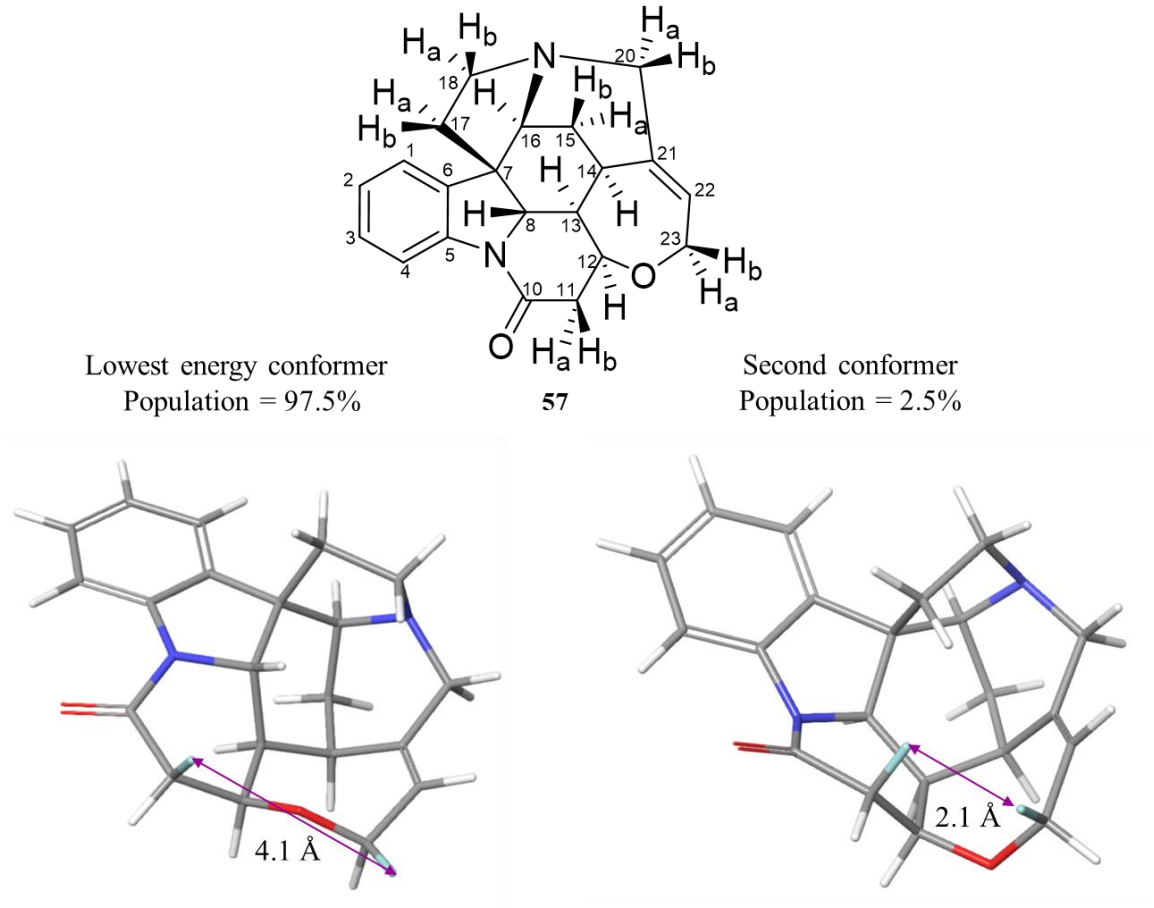


Figure 18. The structure of strychnine **57** and the two conformers found by DFT calculations. The clear conformational change of the seven membered ring is shown giving rise to the shorter H_{11b}-H_{23b} distance.

NMR spectroscopy is an incredibly valuable method for studying molecular structure and conformation in solution as demonstrated in this section, although the methods described are not an exhaustive list of methods available to evaluate molecular conformation/configuration. However, NMR is not without its limitations and the elucidation of conformation for highly flexible molecules by NMR spectroscopy alone is extremely challenging, due to a large number of conformers interconverting on the NMR timescale. Fortunately, the analysis of molecular conformation by NMR spectroscopy can be supported by molecular mechanics (MM) and quantum mechanics (QM) calculations, enabling the elucidation of 3D structure for highly flexible molecules.[110] This will be discussed in the next section.

1.3.2 Aiding Structure Elucidation using Molecular Mechanical and Quantum Mechanical Calculations

For flexible molecules, using NMR to elucidate 3D molecular conformation is challenging because a large number of conformers contribute to the observed NMR parameters (δ, J and interproton distances) making it extremely difficult to estimate the populations of all contributing conformers. For cases such as this, it is possible to use MM and QM calculations to calculate conformer populations, geometries and Boltzmann averaged NMR parameters. These can be directly compared with the experimental NMR parameters and the quality of the fit can be assessed, helping to elucidate conformer populations and ultimately 3D molecular structure.

The first, and arguably most important, stage of this process involves the exploration of conformational space to identify the global and local minima on the potential energy surface. If the conformational space is inadequately sampled, then a poor correlation between experimental and calculated NMR will be observed. For molecules with a small number of rotatable bonds, exploring conformational landscape is relatively straightforward. However, as the number of conformational degrees of freedom increases so does the number of conformers, exponentially.[111]

Different computational methods are available for exploring conformational space, with vastly differing levels of computational intensiveness and accuracy. Very sophisticated *ab initio* QM methods exist and can provide highly accurate conformer geometries and energies.[112] However, these are very time demanding calculations. [113] Alternative methods, employ MM and include molecular dynamics (MD) simulations or Monte Carlo (MC) sampling.[114] MM methods treat the molecule as a 'ball and spring' and calculates the strain energy of a molecule as a sum of independent terms, described by **Equation 4**.[115] By simplifying a molecule to just atoms and bonds, and not including nuclei and electrons, the computational time required to calculate conformer energies is greatly reduced. These terms are parameterised by a force field. Examples include MMFF[116], MMFFs, MM2[117], MM3[118], OPLS_2001.[119] The choice of force field is largely down to the molecule in question.

$$E_{total} = E_{stretch} + E_{bend} + E_{torsion} + E_{coulombic} + E_{Van\ Der\ Waals}$$

Equation 4

MM conformational search methods fall into one of two categories, stochastic or systematic. Systematic methods involve incrementing either cartesian xyz coordinates or torsional angles by a predetermined amount. Although these methods are exhaustive, the exponential increase in the conformational space that is searched with increasing number of rotatable bonds, limits their application to rigid molecules. Stochastic methods, such as MC conformational searching, however are much faster and a more favourable choice for flexible molecules. Stochastic methods generate conformers by distorting the

preceding conformer randomly by a small change. The different conformational search methods are compared in **Table 2**.

Table 2. A comparison of conformational search methods, highlighting their advantages and disadvantages

	Quantum Mechanics (QM)	Molecular Mechanics (MM)
Advantages	- Higher Accuracy - Better at identifying low energy conformers as it is more likely to correctly identify local and global minima	-MD provides information about a system in time (dynamic information) -MC methods are considerably less time demanding -MC methods are very effective at crossing high energy barriers due to the stochastic nature of the search -MD is very effective if solvation is required as explicit solvent models are routinely employed
Disadvantages	- Time Demanding - Unsuitable for high-throughput analysis -No dynamic information	-Lower accuracy of conformer energies and geometries -MD methods remain time demanding for long timescale simulations -MD methods can poorly sample conformational space if large energy barriers to torsional rotation are present

The Monte Carlo Multiple Minimum (MCMM) is a stochastic method routinely used in the conformational search of flexible molecules.[120] During an MCMM conformational search, the starting conformer is distorted randomly based on either cartesian xyz coordinates or torsional angles. The distorted conformer is then minimized, and the resulting conformation is compared to those previously generated. The structure is either retained as a new conformation or it is eliminated if it is either redundant to a previously found conformation or if it is too high in energy. The process is repeated until the number of steps, defined by the user, has been reached (**Figure 19**). Since the MCMM method makes random changes to the structure the conformational space is usually well sampled, provided that enough iterations are defined by the user. To ensure that the calculation finds all the possible conformations a large number of steps are typically required.

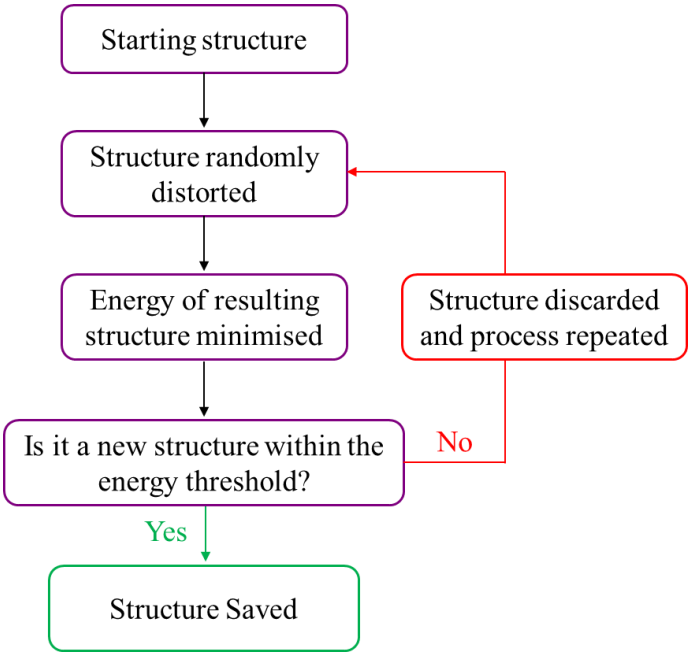


Figure 19. The general workflow followed in a stochastic conformational search such as MCMM.

The energies of the conformers estimated using MM methods (**Equation 4**) are low in accuracy due to the approximation that molecules can be treated as a 'ball and spring'. Furthermore, using MCMM methods, the zero-point energy correction, entropy and solvation are not typically taken into consideration hence reducing the accuracy further. This can be addressed with the use of QM methods such as Density Functional Theory (DFT). Using DFT calculations it is possible to calculate optimised conformer geometries, the Gibbs free energies of the conformers and the NMR parameters including magnetic shielding tensors, which can be converted to chemical shift, and scalar coupling constants.

QM methods, such as DFT, differ from MM methods by approximating the Schrödinger equation and treating the molecules as collections of nuclei and electrons. In DFT, the ground state energy of a system is approximated as a sum of the kinetic energy, the electronnuclear interaction energy, the Coulomb energy and the exchange-correlation energy. These terms all depend on electron density and thus the ground state energy can be summarised as a function of electron density.[121,122] The exact form of the exchange-correlation energy is unknown[123]; however, it can be approximated using one of the many functionals developed to address this such as B3LYP[124] and mPW1PW91[125].

The electron density of the atomic orbitals is described by basis sets, which describe both the core and valence molecular orbitals. Functions can be added to the basis sets to describe the polarisation and diffusivity of the basis sets.[126,127] The most popular choice of basis sets are typically *split-valence* basis sets, with two or more functions describing the valence orbitals. An example of *split-valence* basis sets are the Pople basis sets, which use primitive Gaussian functions to describe the electron density of the core and valence atomic orbitals.[128] The basic form of a Pople basis set is of the type: N-IJKG where

N represents the number of gaussian functions describing the core orbits and I, J and K are the gaussian functions describing the valence orbitals. With three functions describing the valence orbital, it is known as a triple-ζ basis set, however double- ζ basis sets with only two functions describing the valence orbitals also ssexist.[129] Examples of some Pople basis sets are: 3-21G, 6-31G and 6-311G. The addition of polarisation functions, which describe distortions to the symmetry of polarisation of electron density around the nucleus upon the approach of other nuclei, are denoted as either an asterisk or as a letter denoting the polarisation function added.[130] For example, the addition of polarisation functions to 6-311G results in the following notation: 6-311G** or 6-311G(d,p).

DFT can be used to calculate optimised geometries of structures along with the Gibbs free energy of the resulting coordinates. This information can be used to calculate the Boltzmann populations of a set of conformers which can be used to Boltzmann average interproton distances across the set of conformers. DFT can also be used to calculate NMR parameters, including magnetic shielding tensors (MST) and scalar coupling constants.[131,132] The Gauge Independent Atomic Orbital (GIAO) method has been widely used to calculate magnetic shielding tensors.[133,134] The GIAO method, which uses basis functions that have an explicit field dependence, calculates the isotropic component of the magnetic shielding tensors. These can then be converted to a chemical shift by calibrating the isotropic shielding tensors to those calculated for tetramethylsilane (TMS) at the same level of theory.[135] DFT can also be used to calculate scalar coupling by parameterisation of the four terms that contribute to spin-spin couplings including, the Fermi Contact (FC), the diamagnetic spin-orbit (DSO), the paramagnetic spin-orbit (PSO) and the spin-dipole (SD). The scalar coupling constants and chemical shifts calculated for each conformer are averaged using their calculated Boltzmann population, to obtain Boltzmann averaged NMR parameters which can be directly compared to the experimental NMR parameters.

1.3.3 The Application of the Hybrid NMR Spectroscopy and Computational Approach to Solving 3D Molecular Structure

As mentioned previously, the hybrid QM and NMR spectroscopy method has been used throughout the scientific literature to examine 3D molecular structure through the comparison of one or more experimental NMR parameter (*J*, δ or interproton distance) to those calculated using DFT calculations.

One of the earliest examples of using the hybrid NMR spectroscopy and computational approach was reported in 2012 by Butts and Bifulco *et. al.* in the elucidation of the stereochemistry of the natural product Conicasterol F.[108] The work marked the first occurrence of measuring multiple NMR parameters, including chemical shifts and quantitative interproton distances, and comparing to those calculated using DFT calculations. Prior to this, the hybrid QM and NMR spectroscopy approach,

pioneered by Bifulco, involved the comparison of only one NMR parameter and was restricted to using ^{1}H or ^{13}C chemical shifts or scalar coupling constants.[86,110,136,137] The workflow proposed by Butts and Bifulco *et. al.* is shown in **Figure 20**. This workflow, although often tweaked on a case by case basis, has formed the general protocol adopted in the hybrid QM and NMR spectroscopy approach towards structural elucidation of natural products.

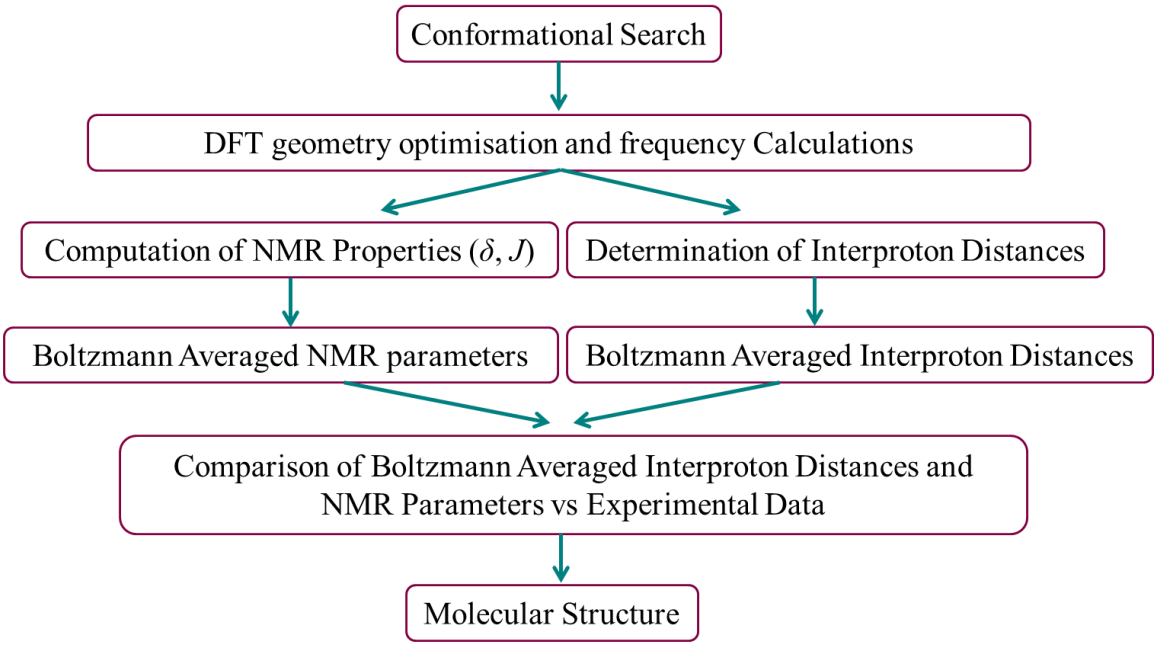
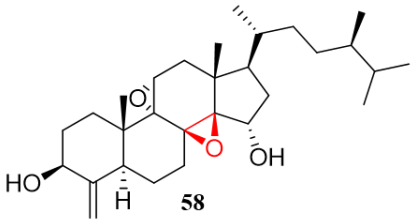


Figure 20. The general workflow employed by Butts and Bifulco *et. al.* in the hybrid QM and NMR spectroscopy approach to solve the relative stereochemistry of Conicasterol F.[108]

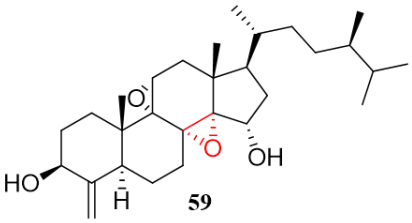
The hybrid QM and NMR spectroscopy approach was used to solve the ambiguity over the relative stereochemistry of the epoxy ring. The calculated ^{13}C chemical shifts and interproton distances, derived from 1D-ROE, for two diastereomers of Conicasterol F were compared to those obtained experimentally. DFT calculations, using the mPW1PW91 functional and 6-31G(d) basis set, were used to calculate δ_c and interproton distances. Calculated δ_c and interproton distances for each conformation were Boltzmann averaged against their calculated Boltzmann populations.

They first examined the comparison of interproton distances and found that diastereomer **59**, as opposed to **58**, clearly gave a better fit between the experimental data of the natural product and calculated data, suggesting that this is the correct diastereomer of Conicasterol F (**Table 3**). To support their hypothesis, calculated δ_c were compared to the experimental δ_c and a much lower MAD was also observed for diastereomer **59**. Confirming that the correct diastereomer of Conicasterol F is **59**, supported by two separate analyses. Comparing multiple NMR parameters is especially useful in cases when the discrimination between diastereomers is not high enough for one of the NMR parameters.

Table 3. The MAD and SD obtained in the comparison of experimental interproton distances and ^{13}C chemical shifts to the calculated data for two diastereomers of Conicasterol F.[108]



58



59

	MAD	SD	MAD	SD
Interproton Distances	7.8%	5.9%	3.0%	2.6%
δ_c ppm	3.7	-	0.8	-

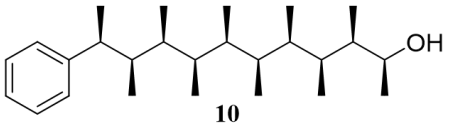
The first example discussed involved the application of the hybrid NMR spectroscopy and computational approach to solve the relative stereochemistry of a rigid steroid, with conformational search calculations identifying a single conformation of the steroid ring system.[108] Performing this analysis on very rigid systems such as this is relatively straightforward process. However, it has been shown by Aggarwal *et. al* that this methodology can be applied to highly flexible open chain compounds with similar levels of accuracy. The conformation of the molecules synthesised by Aggarwal *et. al.* **54**, **55** and **56**, discussed in section 1.2, were explored by the hybrid QM and NMR spectroscopy approach, presenting one of the first examples of this method being applied to a molecule with such a large number of rotatable bonds.[27] The terminal boronic ester was first oxidised to the corresponding phenol before the conformational analysis.

Two of the three molecules synthesised (**10** and **11**) by Aggarwal *et. al.* should display a conformational bias in solution, due to the avoidance of destabilising *syn* pentane interactions. X-ray crystallography was initially used to confirm the expected conformational bias of the molecules however, the conformation of a molecule in the solid state can be influenced by crystal packing and thus the hybrid QM and NMR spectroscopy approach was used to examine the conformational bias in solution. NMR spectroscopy was used to measure quantitative interproton distances, using the PANIC method, as described previously. In addition, H-H scalar coupling constants were either measured directly from the multiplets in the 1D-^1H spectrum or from spectrum simulation using SpinSimulation in MestreNova. Finally, H-C scalar couplings were measured using EXSIDE. The experimental NMR parameters measured were then compared to those calculated using DFT calculations (mPW1PW91/6-311 G(d,p)). The mPW1PW91, a Generalised Gradient Approximation (GGA) functional that improves the approximation of total energy by considering the gradient of the electron density, has been shown by Bifulco *et.al.* to be the best functional to use when calculating chemical shifts using the GIAO method.[138] This provided refined conformer geometries, accurate estimates of conformer energies and

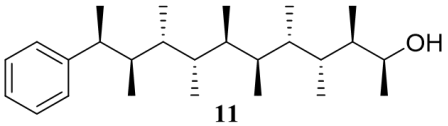
predicted *J* couplings. The scalar coupling constants and interproton distances were Boltzmann averaged against their calculated conformer energies to provided Boltzmann averaged NMR parameters that could be compared to the experimental NMR data.

The DFT calculations predicted that for the alternating *syn-anti* diastereomer **11** >95% of conformers were predicted to be linear. Similarly, for the all-*syn* diastereomer **10** >74% of conformers were predicted to be helical (**Table 4**). For the all-*anti* diastereomer (**56**) a MM conformational search revealed no conformational bias, and thus no further computational or NMR work was performed. For both **10** and **11** an excellent correlation between the experimental and calculated NMR parameters were observed with a 1 Hz MAD in the comparison of scalar coupling constants and a 5% error in the comparison of interproton distances. Although these deviations are marginally larger than those observed previously on very rigid structures, they are in line with those obtained for flexible molecules. The excellent correlation observed between experimental and calculated NMR parameters, confirms that the conformational bias predicted computationally is present in solution.

Table 4. The MADs obtained between experimental and computational NMR datasets for the two diastereomers **10** and **11**.



10



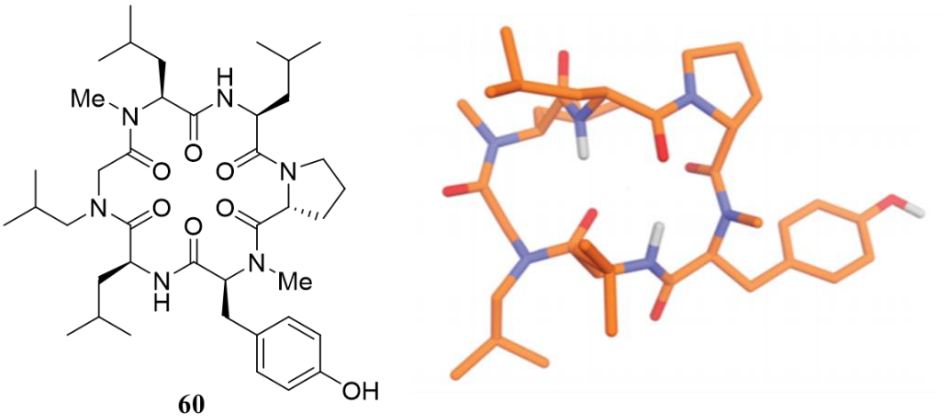
11

	MAD	**MAD**
Interproton Distances	5.8%	4.2%
$^{n}J_{HH}$ Hz	0.71 Hz	0.66 Hz
$^{n}J_{CH}$ Hz	0.50 Hz	0.42 Hz

The previous two examples involved the comparison of multiple NMR parameters. However, it is possible to use only one NMR parameter in the hybrid QM and NMR spectroscopy approach. Commonly, ^{1}H or ^{13}C chemical shifts are used as these are experimentally very simple to measure. Additionally, by removing the calculation of scalar couplings from the NMR calculations, the computational time required is greatly reduced. This allows for a comparison between experimental and calculated data to be obtained much faster although often with less discrimination.

The comparison of experimental and calculated chemical shifts has pioneered by the research group of Dean Tantillo, who have reported the use of this methodology on a structurally and flexibly diverse range of molecules.[139] Recently, they have shown that this methodology can be performed on cyclic peptides of up to 6 residues (**Figure 21**, **60**).[140] Analysing the structure and conformation of cyclic peptides is particularly challenging due to the unusual chemical environments that are often observed, caused by shielding or deshielding from ring current effects. This results in chemical shifts that are challenging to predict by QM methods. Tantillo *et. al.* report the use of MM conformational search

calculations using MCMM search methods, followed by QM calculations to predict cyclic peptide ^{1}H and ^{13}C chemical shifts. When compared to the experimental ^{1}H and ^{13}C chemical shifts, good MADs of 2.3-3.1 ppm for ^{13}C and 0.27-0.3 ppm for ^{1}H were obtained. However, Tantillo *et. al.* have since shown that the MAD can be significantly improved to <2 ppm for ^{13}C and <0.2 ppm for ^{1}H when using MD to explore conformational space.[141] They showed that using high temperature MD simulations, allowing for *cis-trans* rotamerisation of the amide bonds to occur, captured more subtle conformational motion than MM resulting in a better description of the conformational landscape.



60

Figure 21. An example of a cyclic peptide studied by Tantillo *et. al*. This figure was reproduced from reference [141]

The previous examples have employed the use of statistical analyses to compare experimental and calculated data, namely MAD and SD. It is also possible to use probability-based methods to quantify the quality of fit between the two datasets. One example that has been widely adopted is DP4, reported by Goodman *et. al.* in 2010.[135] Probability based methods usually offer a greater discrimination between diastereomers than statistical analyses.[142] DP4 involves the calculation of ^{1}H and ^{13}C chemical shifts by QM calculations followed by an empirical scaling of the predicted chemical shifts to remove any systematic errors. An error between the experimental shifts and the calculated scaled shifts is obtained ($\delta_{calc} - \delta_{exp}$) for each shift. Assuming that each error follows a normal or *t* distribution, then the probability of encountering each error in a correct structure is calculated. The probabilities for each signal are multiplied to give the absolute probability that the structure is correct. DP4 has been shown to provide high levels of discrimination between diastereomers in the stereochemical assignment of many natural products however, it only considered the comparison of chemical shifts and there are many situations where chemical shifts alone are not capable at confidently assigning stereochemistry, and in these situations an approach that considers multiple NMR parameters is required.

1.4 The Inhibition of Protein–Protein Interactions

As discussed previously, conformationally preorganising a drug candidate towards the bound conformation is particularly important in the inhibition of protein–protein interactions (PPIs). This section will discuss what PPIs are, why they are such important therapeutic targets, the challenges associated with inhibiting PPIs and the breakthrough strategy that turned PPIs into 'druggable' targets.

1.4.1 What are Protein–Protein Interactions?

Proteins are responsible for the structure, function and regulation of our bodies tissues and organs and are integral in almost all cellular processes, from catalysing cellular reactions to transmitting the information needed for DNA replication. However, proteins rarely act alone and many of these functions are dependent on proteins forming specific interactions with other proteins, forming what is known as protein–protein interactions (PPIs).[143] PPIs are specific, they are not a chance or accidental interaction with another protein.[144] The complex network of interacting proteins is known as the interactome and the human interactome is estimated to include >650,000 distinct pair-wise interactions, estimated using a statistical analysis of the wealth protein interaction data available in the literature.[145] Mapping this interactome and understanding how the hundreds of thousands of interactions regulate cellular function is by no means a simple challenge.

Since PPIs play such a crucial role in the regulation of so many cellular functions, they also possess the ability to cause numerous diseases if they become aberrant. Indeed, a number of diseases including cancer, Alzheimer's and HIV, can be attributed to misregulated PPIs.[146] Despite this therapeutic need, very few strategies to target these misregulated PPIs currently exist. This is due to the nature of the binding interface that typically exists between PPIs, which is not one that can be considered to be easily 'druggable'.[147] Typically, the design of a competitive inhibitor of protein function will involve the design of a small molecule that fits into a well-defined pocket on the protein of interest. In addition to high shape complementarity between the ligand and protein, often dominant interactions such as hydrogen bonding, salt-bridges or covalent interactions exist. This is seen in the 'lock-and-key' mechanism for the inhibition of enzyme-substrate complexes with a small molecule and has been hugely successful in the pharmaceutical industry.

The binding interface between proteins in a PPI however does not involve one protein binding to a well-defined small pocket on another protein. Instead, it involves very large binding interfaces of >1000 $Å^2$ between the two proteins (**Figure 22**).[148] Additionally, the binding interface is shallow, featureless and non-contiguous.[149] The binding between proteins in a PPI is also typically dominated by the hydrophobic effect and very few specific interactions, such as hydrogen bonds, exist between the two proteins.[150] Obtaining selectivity in in the inhibition of PPIs is therefore a formidable challenge as

there are often no specific interactions between the two interacting proteins. It is for these reasons that PPIs were originally considered to be 'undruggable' targets because it was extremely challenging to design a small molecule that could cover the large surface areas and complement the poorly defined binding interface across a relatively flat surface.[145]

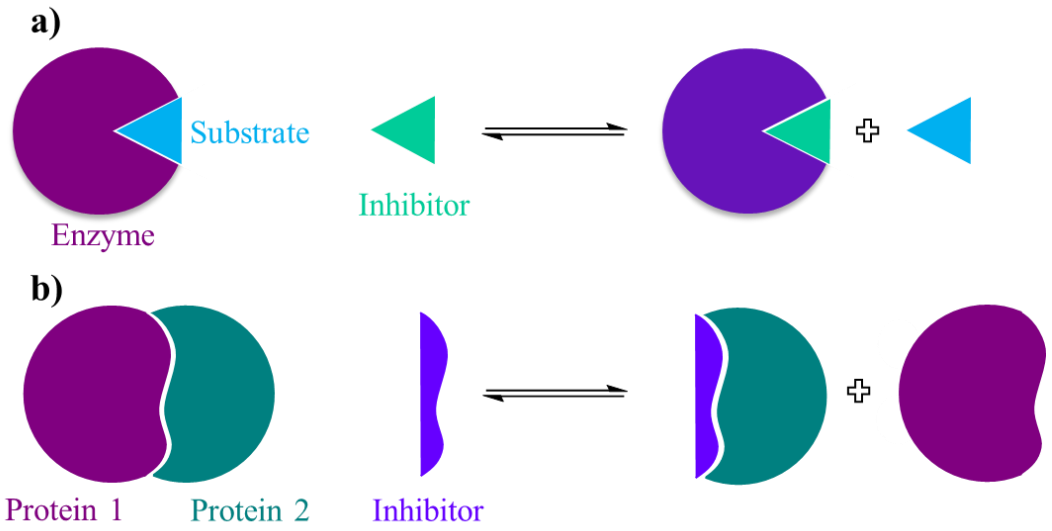


Figure 22. The inhibition of **a**) an enzyme-substrate interaction and **b**) a protein–protein interaction

1.4.2 How can Protein–Protein Interactions be Disrupted?

Perhaps the obvious choice in the inhibition of PPIs would be the use of peptide based therapeutics. Peptides offer a logical strategy to protein inhibition since the natural protein sequence can be mimicked by the peptide.[151] This would offer the advantages of greater selectivity and specificity over a small molecule. Additional benefits of peptide based therapeutics include their ease of synthesis and better pharmacokinetic properties and less off target side effects.[152] However, peptide based therapeutics exhibit poor membrane permeability and rapid proteolytic digestion. Additionally, they often show poor conformational stability which decreases affinity to the target as discussed previously.[50,153] Although strategies have been developed to address these issues, such as peptide stapling, peptides remain the less favoured choice when compared to small molecules. However, small molecules are also not well suited for the inhibition of PPIs, due to the nature of the binding interface. That was, until the realisation that the binding affinity between PPIs is typically dominated by a few key residues.

In 1995, Wells and Clackson reported that although PPIs take place over a large interface, the interactions that drive formation of the PPI are focussed on a few key amino acids, termed hot-spot residues (**Figure 23**).[154] Using alanine mutagenesis studies, whereby each residue is sequentially mutated to alanine and the effect on the affinity to its protein partner is computationally investigated, it was revealed that the free energy of binding was dominated by a few amino acids.[155] Typically, the

hot-spot residues constitute less than half of the binding interface of a PPI and are located at the central region of the binding interface.[156] Additionally, the hot-spot groups are often found in tightly packed clusters and interact cooperatively.[157] The hot-spot approach to PPI inhibition proposed by Wells and Clackson was ground-breaking for the development of small molecule PPI inhibitors and it suggests that a small molecule that simply mimics these hot-spot as opposed to the entire binding interface, will be sufficient for competitive inhibition. Indeed, the hot-spot approach kickstarted the development of small molecule inhibitors of PPIs and now a number of therapeutic molecules are present at various stages in the drug discovery pipeline.[158,159]

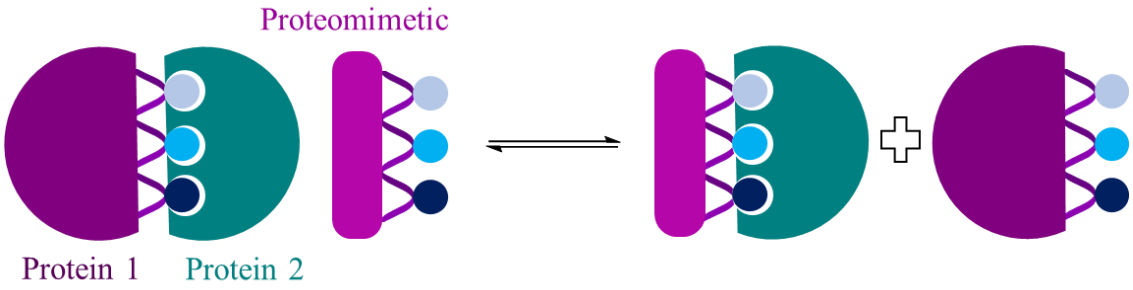


Figure 23. The hot-spot approach to PPI inhibition proposed by Wells and Clackson. The three hot-spot groups are represented as blue circles. They bind to a groove or pocket on a second protein.[160]

1.5 The p53-Mdm2 Protein–Protein Interaction

Across all 650,000 PPIs that exist in the interactome, the p53-Mdm2 PPI has captured scientist's attention and become one of the most studied in the scientific literature. This next section will discuss the p53-Mdm2 pathway, the mechanisms by which this PPI becomes aberrant and the current therapeutic pathways aimed at treating this PPI.

1.5.1 The p53-Mdm2 Pathway

p53, a tumour suppressor protein, is responsible for protecting cells from malignant transformation giving it the rightly deserved nickname the 'guardian of the genome'.[161,162] p53 is a powerful transcription factor that regulates the cellular pathways responsible for, although not limited to, DNA repair, cell cycle arrest, apoptosis, angiogenesis and senescence.[163],[164]

Human p53 contains 393 amino acids and consists of different structural and functional domains. It is biologically active as a homotetramer. The first 42 amino acids at the N terminus form the transactivation domain, responsible for transcriptional activation through the interaction with various transcription factors.[163] Following this is a proline rich region from 61-94, which plays a role in a regulatory functions including p53-mediated apoptosis.[165,166] The core of p53, residues 94-292, is known as the DNA binding domain and is responsible for sequence specific binding to double stranded

DNA.[167] Residues 324-355 make up the C-terminal tetramerization domain, responsible for oligomerization of p53.[168,169] Finally, a regulatory domain from residues 363-393 is present at the C-terminus. Residues in this region undergo posttranslational modifications including phosphorylation, methylation and ubiquitination, promoting structural changes to the central region of p53, controlling the sequence specific binding of p53 to DNA.[170],[171]

The N-terminal transactivation domain is also the location of binding of Mdm2, p53's negative regulator.[172] Mdm2 is an oncogene whose primary responsibility is to modulate p53's tumour suppressor activity through an autoregulatory feedback loop. Mdm2 has been shown to make extensive contacts with several hydrophobic residues, namely F19, L22, W23 and L26, on p53. Two of these amino acids, L22 and W23, are also key to the interaction of p53 with transcription factors.[173,174] Thus, when bound to Mdm2, p53 loses its transcriptional activation activity.

The regulation of the Mdm2-p53 autoregulatory feedback loop is critical in protecting cells from damage. The role of the p53-Mdm2 autoregulatory loop is to maintain low cellular concentration of p53 in healthy cells and to increase p53 concentration on response to cellular damage.[175] Mdm2 regulates the activity of p53 through different mechanisms. As mentioned previously, Mdm2 blocks the transactivation domain of p53, preventing p53-mediated transcription.[176] Mdm2 also promotes p53 degradation through ubiquitination as it is an E3 ligase protein. Finally, Mdm2 binding induces the nuclear export of p53. If p53 is not present in the nucleus it can no longer bind to DNA. Through these three mechanisms, the levels of cellular p53 are tightly controlled by Mdm2 and p53 is present at low concentrations in healthy cells.[177,178] The autoregulatory negative feedback loop is completed as Mdm2 is itself a product of p53-mediated transcription (**Figure 24**).[179]

In response to cellular stress however, this feedback loop is disrupted allowing p53 to accumulate and bind to DNA. Several pathways resulting in the activation of p53 through disruption of the p53-Mdm2 feedback loop exist. For example, following DNA damage, specific residues in the N-terminal transactivation domain are phosphorylated (S15, T18 and S20) disrupting the interaction between p53 and Mdm2.[180] When p53 is unregulated by Mdm2 it can enter the nucleus, bind to DNA and promote the transcription of genes responsible for protecting the cell. Additionally, p53 can now promote the transcription of Mdm2, increasing cellular concentrations of Mdm2.[181] This causes a subsequent decrease in the concentration of p53 back to the baseline levels. This autoregulatory feedback loop is critical for the survival of cells however, numerous mechanisms exist that prevent the correct functioning of this feedback loop, resulting in malignant cells going unregulated.[182]

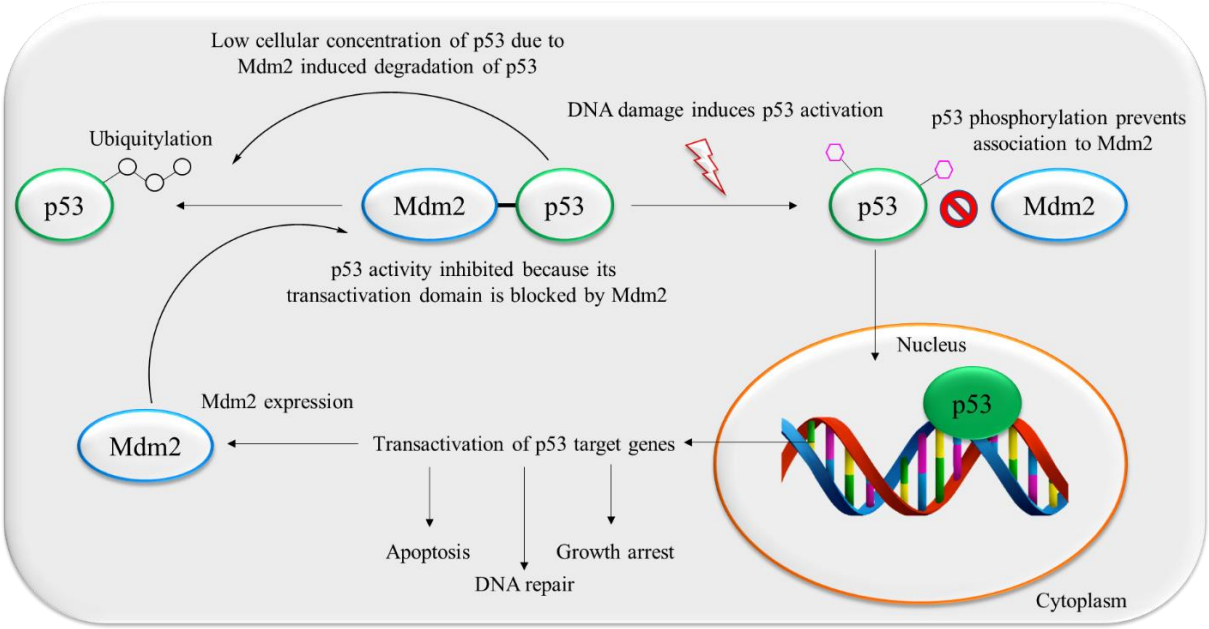


Figure 24. The autoregulatory feedback loop of p53 and Mdm2.

1.5.2 Disruption of the p53-Mdm2 PPI through Mutation of p53

The correct functioning of the p53-Mdm2 autoregulatory feedback loop is critical for the survival of healthy cells, and interruptions to this loop directly contribute to tumour growth. In fact p53 is the most frequently inactivated protein in human cancers with over 50% of diagnoses linked to misregulation of the autoregulatory loop between p53 and Mdm2.[183] Due to this, restoration of the p53-Mdm2 feedback loop is the key aim of many cancer treatments and it is capable of halting tumour growth.

The most frequent interruption to this loop is due to the mutation of p53, which causes the loss of p53 activity.[184] 95% of the p53 mutations that have been reported are point mutations located in the centre of the sequence-specific DNA binding domain and influence either the conformation of p53 or the residues responsible for making contacts with DNA.[185],[186] Despite mutations in p53 being responsible for approximately half of all new cancer diagnoses, there are very few drugs in the pipeline designed to restore the function of mutated p53.

A strategy that has showed signs of promise is the design of a small molecule that can bind to either a transient or long-lived pocket on mutated p53 for long enough to restore the conformation of p53 to its native conformation allowing it to bind to DNA.[187] PRIMA-1 (**61**), or the more potent methylated version, PRIMA-1^{MET} (**62**) (or APR-246) developed by Aprea AB, is currently the most successful small molecule drug capable of restoring mutated p53 function. PRIMA-1^{MET} is now in stage II clinical trials for women exhibiting ovarian cancer, which is very commonly caused by p53 mutations.[188] The exact mechanism of PRIMA-1^{MET} is not yet fully understood however it was discovered that PRIMA-1 and PRIMA-1^{MET} are decomposed *in vivo* to methylene quinuclidinone (MQ, **63**).[189] It is believed that MQ

reacts covalently with thiol residues on mutated p53 and restores the conformation of mutated p53 to wild-type p53 (**Figure 25**).[190]

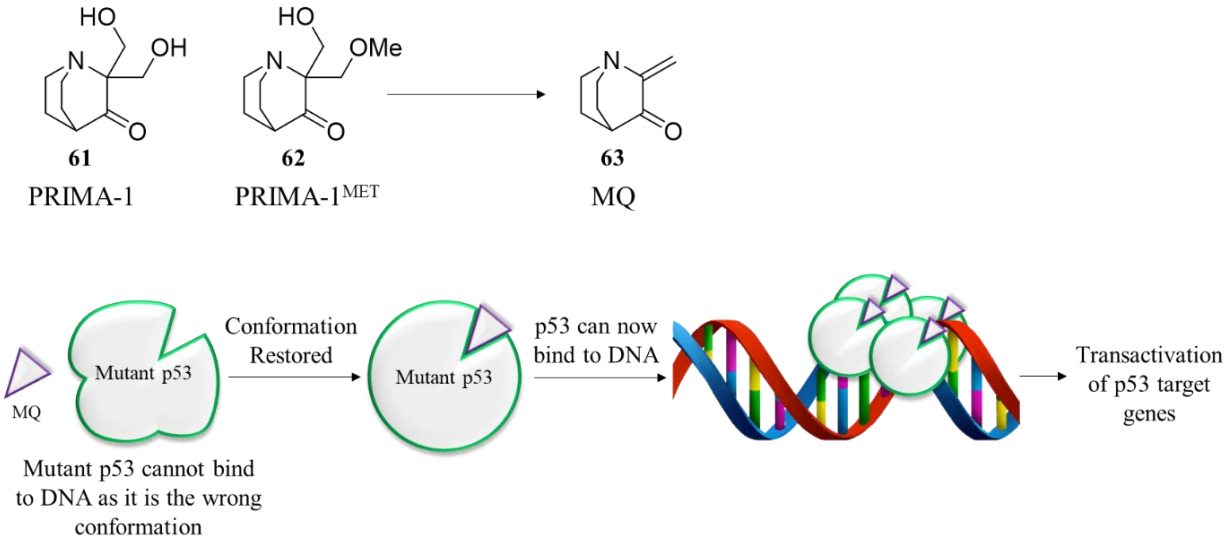


Figure 25. A schematic of the suggested mechanism of PRIMA-1^{MET} towards p53 reactivation.

The proposed mechanism of action of PRIMA-1^{MET} was confirmed by extensive computer simulations, performed by Amaro *et. al.* Using all atom explicit solvent molecular dynamics simulations, they were able to examine how p53 mutations destabilise the conformation of p53 and prevent it from binding to DNA.[191] The simulations confirmed that in mutant p53 a small pocket transiently forms in the core of the protein (**Figure 26**). This cleft is large enough to capture a small molecule, supporting the theory that a small molecule can be used to bind to mutated p53. Present within the cleft are three cysteine residues, all of which are possible targets for alkylation with MQ. The simulations performed by Amaro *et. al.* confirmed that MQ alkylates Cys124 and restores the conformation of mutant p53 to wild-type p53. The irreversible alkylation of a cysteine residue by MQ is a significant disadvantage of the molecule, since once it is bound to the transient cleft it remains there and increases the risk of toxic side effects. Additionally, cysteine residues are located on numerous other proteins, not just p53, although it has been proposed that the shape complementarity between the transient cleft of mutated p53 and MQ promotes alkylation of this site over other cysteine residues. A more desirable strategy is a small molecule that can temporarily lodge into the cleft and remain there for long enough to restore p53 function, and then eventually fall off, reducing the risk of side effects. Amaro *et. al.* screened a library of over 1 million compounds against mutated p53 in the hope that a small molecule that met these requirements could be identified. Several hundred compounds were found to bind to the cleft transiently and restore p53 function. Those that succeeded in cellular testing are now being licensed by Actavalon.

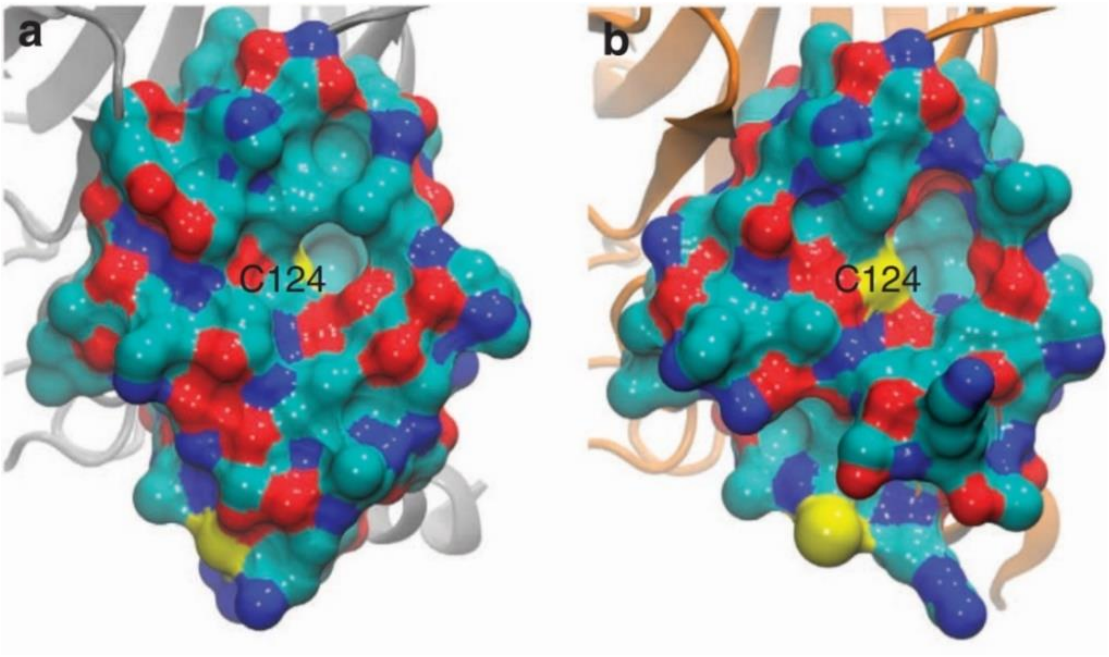


Figure 26. Visualisation of the DNA binding domain of p53. **a)** The crystal structure of wild-type p53 shows that Cys124 is obstructed. **b**) MD simulations of mutated p53 reveal a transiently open pocket that exposes Cys124. This pocket can be targeted by small molecules. This figure has been directly reproduced from Amaro *et. al.*[191]

1.5.3 Disruption of the p53-Mdm2 PPI through Overexpression of Mdm2

The mutation of p53 is not the only mechanism by which the misregulation of this PPI can result in malignancy. It has also been found that in ~30% of human cancers that Mdm2 is overexpressed and as a result of the autoregulatory feedback loop, levels of cellular p53 are decreased.[192] This prevents activation of the p53 pathway upon cellular stress, allowing malignant cells to go unchecked.

The therapeutic targeting of the p53-Mdm2 pathway has primarily focussed on solving the issue of Mdm2 overexpression through the design of small molecules that can inhibit the p53-Mdm2 PPI by competitive inhibition of Mdm2 and reactivation of the p53 pathway (**Figure 27**).[193] As discussed previously, targeting a PPI with a small molecule is challenging due to the large interfacial areas and featureless binding sites. However, the hot-spot approach described by Wells provides a strategy for drugging PPIs with a small molecule. p53 binds to a hydrophobic pocket on Mdm2 with the binding affinity being dominated by a Phenylalanine residue (Phe19), a Tryptophan residue (Trp23) and a Leucine residue (Leu26).[194] The three residues are located on one face of an α-helical region of p53. This poses the possibility of inhibition of this PPI with a small molecule. If a small molecule that binds competitively to Mdm2 can be developed, it is possible that it could be used in the treatment of cancers caused by the overexpression of Mdm2 and reactivate p53, thereby protecting the cell. The next section will discuss the significant developments that have been made in the search of small molecule p53-Mdm2 inhibitors.

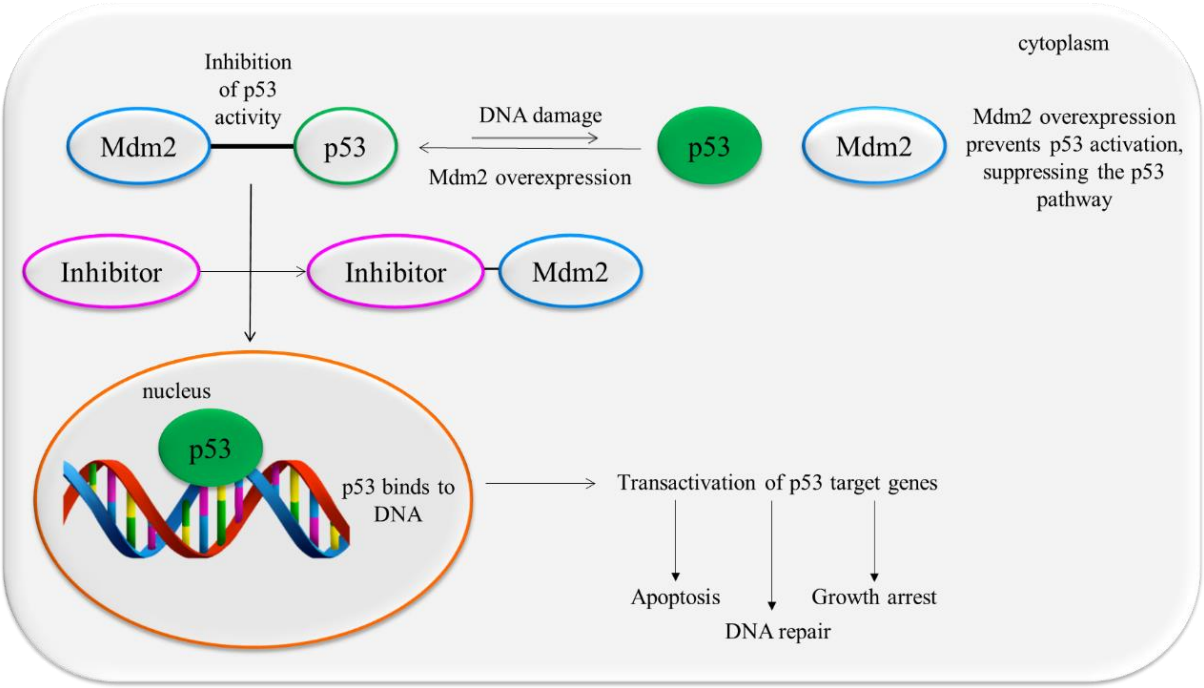


Figure 27. The inhibition of the p53-Mdm2 PPI with a small molecule inhibitor of Mdm2.

1.5.4 Small Molecule Inhibitors of the p53-Mdm2 PPI

The search for small molecules that can disrupt the p53-Mdm2 PPI has not been straightforward, and currently there are no commercially available p53-Mdm2 inhibitors available despite over two decades of intense interest from both academia and industry. The first successful attempt at drugging the p53-Mdm2 PPI was made by a team at Hoffman-La Roche, led by Lyubomir Vassilev.[195] Vassilev *et. al.* identified a family of potent and selective small molecule inhibitors of Mdm2, known as the Nutlins (**64**, **65** and **66**).[195] The Nutlins are a family of tetra-substituted imidazolines that bind competitively to Mdm2 in the p53 binding pocket (**Figure 28**). They have been shown capable of reactivating the p53 pathway in cells overexpressing Mdm2. Nutlin-3 (**66**) is the most potent member and is able to disrupt the p53-Mdm2 PPI with an IC_{50} of 90 nM.

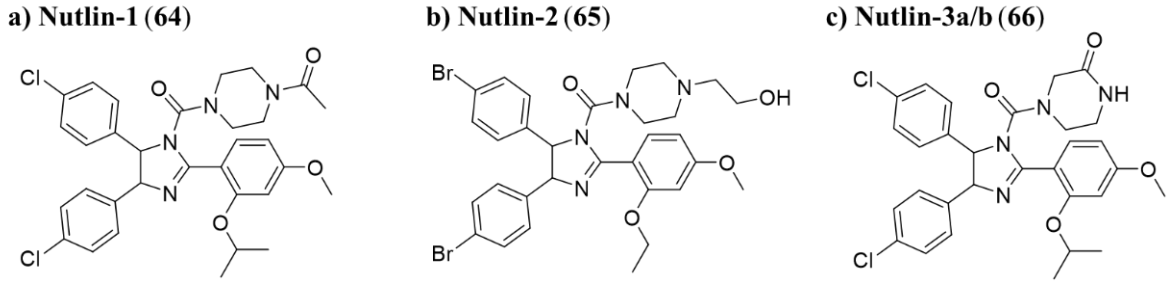


Figure 28. The Nutlins are a potent family of Mdm2 inhibitors. The structures of **a)** Nutlin-1, **b)** Nutlin-2 and **c)** Nutlin-3 are shown.

The binding of the Nutlins to Mdm2 was confirmed using x-ray crystallography, and a crystal structure of nutlin-2 bound to Mdm2 was obtained (PDB ID = 1RV1) (**Figure 29**). The crystal structure revealed that Nutlin-2 binds in a similar fashion to p53, placing three hydrophobic groups into the same position typically occupied by the three hot-spot groups of p53. A 4-bromo-phenyl group occupies the Trp23 pocket, a second 4-bromo-phenyl group is located in the Leu26 pocket and the ethyl ether side chain is directed toward the Phe19 pocket. The hydrogen bond between the indole NH of the Trp23 residue of p53 and the backbone carbonyl of Leu54 on Mdm2 is not maintained however, the excellent binding affinity observed suggests that this interaction is not essential for competitive inhibition. It is interesting to note that a non-aromatic group of Nutlin-2 has occupied the phenylalanine pocket, whilst an aromatic group of Nutlin-2 has occupied the leucine pocket. It has been suggested that this is to allow for a π-π stacking interaction between one of the 4-bromo-phenyl groups and a histidine residue on Mdm2 (His96). The crystal structure also revealed that two molecules of Nutlin-2 are initially associated with the surface of Mdm2. It has been proposed that before Nutlin-2 occupies the p53 binding pocket of Mdm2, it first makes contact with a second site and then tumbles into the binding pocket. It has been suggested that the competitive binding of the Nutlins to Mdm2 is due to the presence of the chlorine atom in the Trp23 pocket.[196] When this pocket is occupied with Trp, a deep void inaccessible to the

indole is present. This void is perfectly occupied by a chlorine atom and is a contributing factor to the competitive binding of the Nutlins. The Nutlins feature a conformationally rigid imidazoline core, which orients the binding groups towards the correct pockets. Due to the rigidity of the Nutlins it is likely that the bound conformation reflects the conformation of the free ligand in solution, and no major conformational reorganisation is required.

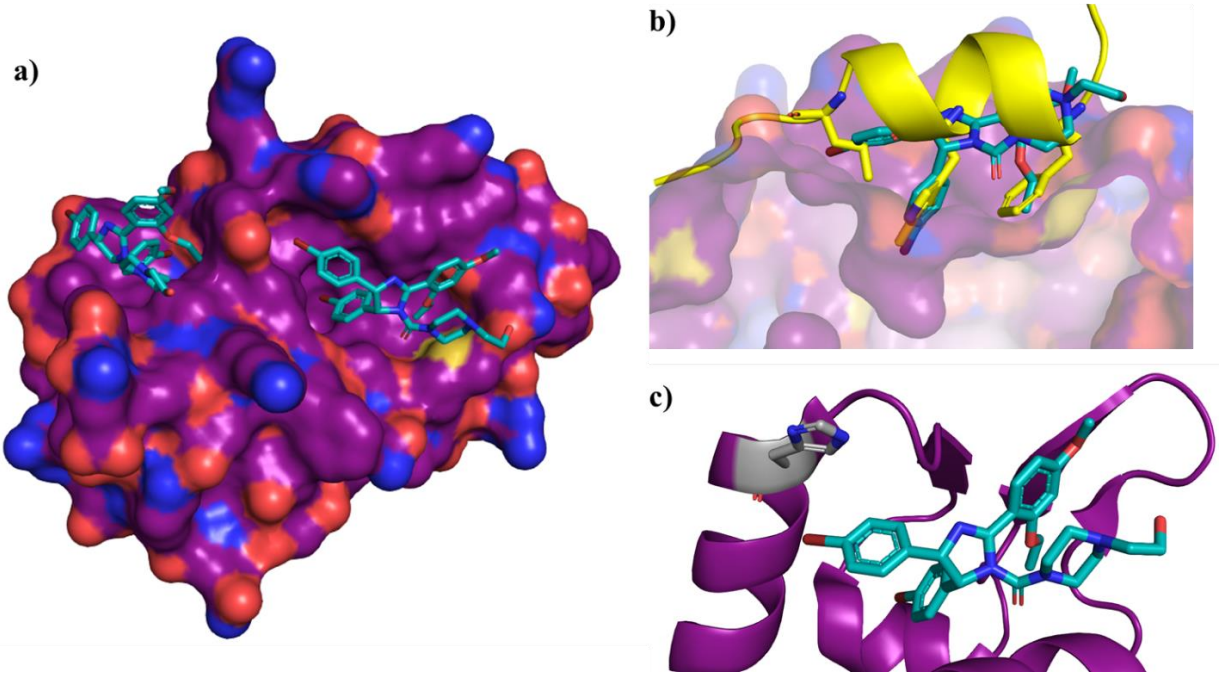


Figure 29. **a)** The two binding sites of Nutlin-2 (**65**) to the p53 binding site of Mdm2. **b)** The overlay of p53 with Nutlin-2 bound to Mdm2. It is clear that the rigid conformation of Nutlin orients three hydrophobic groups into the three pockets of Mdm2 in an analogous fashion to p53. **c)** The proposed π-π interaction between His96 (shown in grey) of Mdm2 and a 4-bromophenyl group of Nutlin-2.PDB ID: 1RV1

The Nutlins showed an impressive ability to selectively bind to the p53 pocket of Mdm2 and reactivate the p53 pathway through inhibition of the interaction between Mdm2 and p53.[196] However, the Nutlins showed less than favourable pharmacokinetics and were metabolically unstable. Roche therefore optimised the structure of the Nutlins to improve their binding affinity, cellular potency, pharmacokinetics and chemical stability.[197] The resulting structure, RG7112 (**67**), retained the imidazoline core with four groups projecting from this central region. RG7112 was the first small molecule inhibitor of Mdm2 to enter clinical trials. The structure of both Nutlin-3a and RG7112 are shown in **Figure 30**. The key difference is the addition of two methyl groups to the imidazoline ring to prevent oxidation of the ring. The isopropyl ether group of Nutlin-3a was replaced with an ethyl ether group as this was assumed to retain good binding with Mdm2 while reducing molecular weight. In Nutlin-3a the methoxy group is prone to metabolic digestion and has been replaced with a *tert*-butyl group to improve metabolic stability. The addition of the sulfonyl group was aimed at improving the pharmacokinetics. A crystal structure of RG7112 reveals that it binds to Mdm2 in a very similar fashion

to that of Nutlin-2. The crystal structure suggests that the π-π interaction between a 4-chlorophenyl group on RG7112 and His96 is better aligned than observed in the crystal structure of Nutlin-2 and Mdm2. RG7112 gave an improved IC_{50} of 11 nM and has been shown to activate wild-type p53 both *in vitro* and *in vivo*, causing the inhibition of cell growth in cancer cell lines.[198]

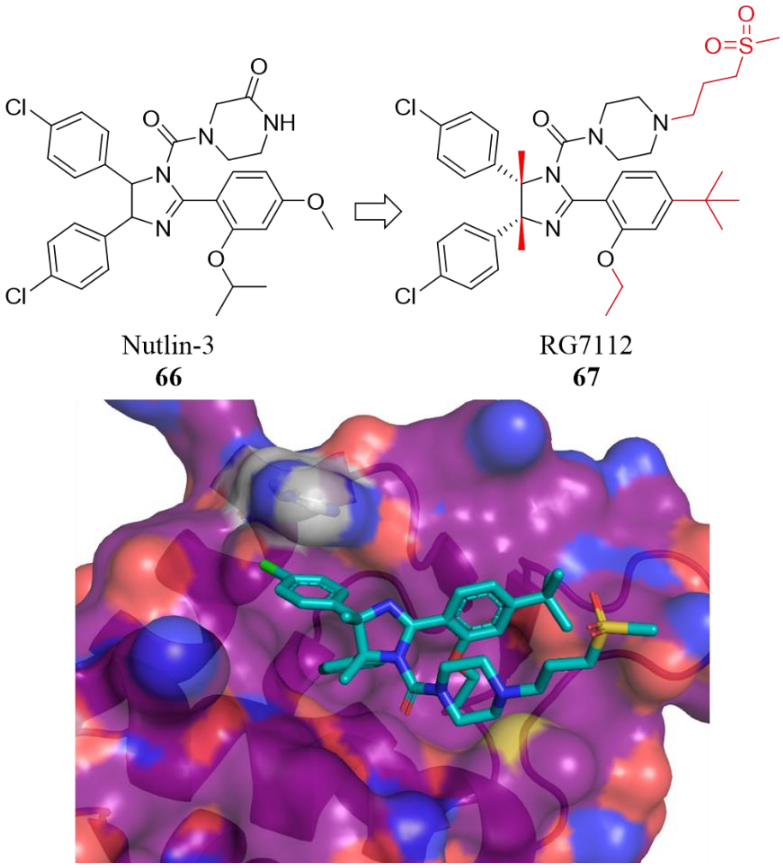


Figure 30. The structural revisions of Nutlin-3a (**66**) to result in RG7112 (**67**). The crystal structure of **67** bound to Mdm2 (PDB: 4IPF) shows that it binds in a similar fashion to the Nutlins. The π-π interaction between the 4-chlorophenyl group and His96 (grey) is better aligned than observed for the Nutlins.

The Nutlin scaffold has inspired many similar small molecule Mdm2 inhibitors that feature a rigid core with four groups projecting from it. One class of structures with many similarities to the Nutlins is the spiro-oxaindoles, first reported by Ding *et. al.*[199] The most potent of the family, MI-219 (**68**), demonstrates good pharmacokinetics and has an IC_{50} of 45 nM. Although the spiro-oxaindoles display good pharmacokinetics and potency, their conformation is not optimal. As discussed previously, it is thought that an aromatic group is preferred in the Leu26 pocket due to a π-π stacking interaction with His96. However, for MI-219 the neopentyl group is occupying the Leu26 pocket. The group at Roche, responsible for many of the previous successes in Mdm2 inhibition, revised the scaffold to switch the orientation of binding and place an aromatic group in the Leu26 pocket (**Figure 31**).[200] The scaffold designed by Roche, known as RG7388 (**69**), is based on a pyrrolidine scaffold.

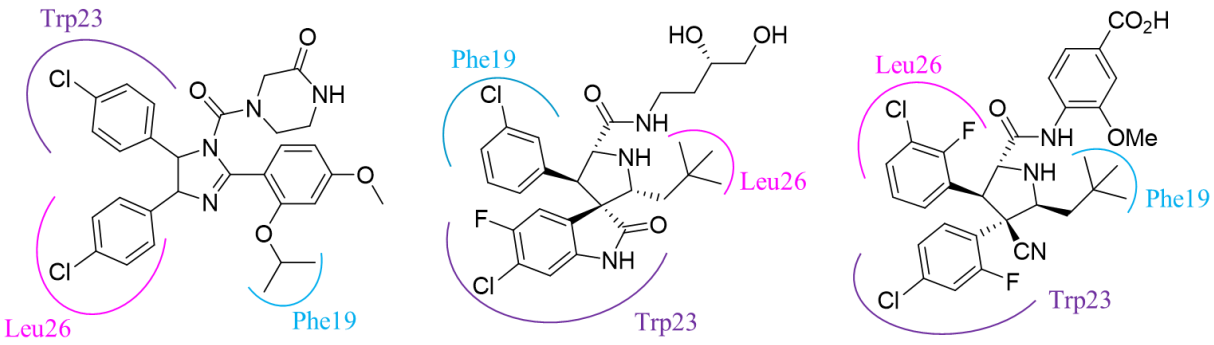
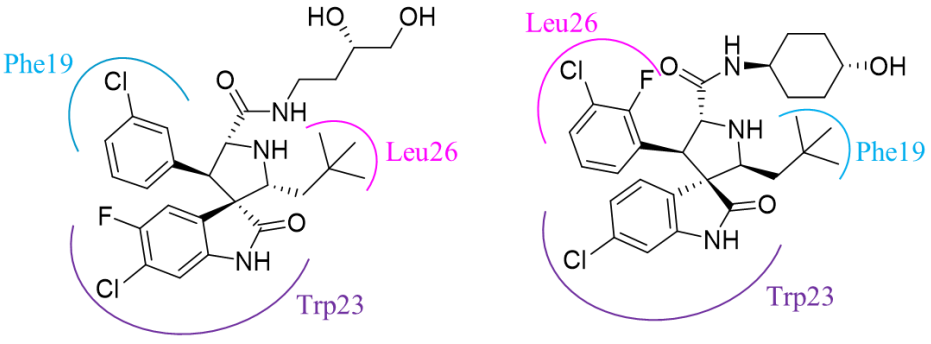


Figure 31. The structure of Nutlin-3a (**66**), RG7338 (**69**) and MI-219 (**68**). The binding pockets that the groups are located in are highlighted.

The group at Roche examined the conformation of the spiro-oxaindole MI-219 and realised that by changing the relationship of the groups on the pyrrolidine ring they could change the orientation of binding of the molecule to place an aromatic group in the Leu26 pocket. For MI-219 the oxindole group and the 3-chlorophenyl group are *cis* to each other whilst the neopentyl group and the oxindole are *trans* to each other. This results in a conformation of MI-219 that places the neopentyl group in the Leu26 pocket. Roche realised that by introducing a *trans* relationship between the two phenyl groups and maintain the *trans* relationship between the 3-chloro-2-phenyl group and the neopentyl group will result in the former being oriented towards the Leu26 pocket.[201] RG7388 is now in clinical trials and shows superior potency and selectivity to both **68** and **66**. RG7388 binds to Mdm2 with an impressive IC_{50} of 6 nM and shows potent tumour inhibition in cancer cell lines containing wild-type p53, moderate clearance ($t_{1/2}$ = 1.6 h) and good metabolic stability.

It is also possible to switch the orientation of binding of the spiro-oxaindoles by changing the stereochemistry of the quaternary centre. The optimisation of MI-219 to incorporate this change, along with a different halogen substitution pattern on both phenyl groups and a conformationally constrained cyclohexanol group gave compound MI-77301 (**70**) (**Figure 32**).[202] MI-77301, has a very impressive K_i of 0.88 nM and has been shown to activate wild-type p53 in cancer cell lines at concentrations as low as 30 nM *in vitro*. MI-77301 is the most potent Mdm2 inhibitor reported to date and has also entered clinical trials.



MI-219
68

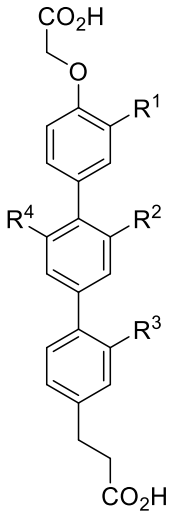
MI-77301
70

Figure 32. The structure of both MI-219 (**68**) and MI-77301 (**69**). The binding pockets that the groups occupy are highlighted.

The majority of small molecule inhibitors of the p53-Mdm2 PPI are based on the original Nutlin scaffold however, there are a number of other structures that bear little to no similarity to the Nutlin scaffold. Following the Nutlins, and the scaffolds that they have inspired, the next most well-known small molecule inhibitor of the p53-Mdm2 PPI are the terphenyls, designed by Hamilton *et. al.*[202] Their importance in the scientific literature is not necessarily due to their inhibitory activity against Mdm2, but rather that they represented one of the first classes of small molecule α-helix mimetics. The importance of small molecule α-helix mimetics and the progress in the field will be discussed in section 1.6, including the conception and development of the terphenyls.

A comprehensive library of terphenyl scaffolds bearing different substituents at the *ortho* position of the three phenyl rings were prepared by Hamilton *et. al.* and their ability to inhibit the p53-Mdm2 PPI was explored using fluorescence polarisation assays and ^{1}H-^{15}N HSQC spectroscopy.[203] The modular synthesis of the terphenyls which sees phenyl groups connected by simple Suzuki cross coupling reactions, allows a broad range and combination of *ortho*-substituents to be screened, helping to identify hot-spot mimetic groups that provide high affinity to Mdm2. A selection of the terphenyls synthesised by Hamilton *et. al.* are shown in **Table 5**. The results obtained confirmed the importance of the *ortho* side chains in establishing good affinity with Mdm2 as removal of one or more side chains resulted in a loss of affinity. Additionally, a comparison between **72** and **73** reveals that the 2',6'-dimethyl substituted terphenyl shows improved affinity over its monomethyl substituted counterpart, likely due to the increased conformational rigidity of the terphenyl backbone imposed by the sterically hindering methyl group preventing rotation around the aryl-aryl bonds. Finally, the results obtained suggest that the naphthyl group is a suitable mimetic of the indole group of tryptophan with strong binding affinities obtained when the naphthyl group is present.

Table 5. A selection of the terphenyl scaffolds designed by Hamilton *et. al.* towards the inhibition of p53-Mdm2.[203]

Structure		R^1	R^2	R^3	R^4	K_i ± S.D. (μM)
	**71**	H	H	H	H	>1000
	72	Bn	Me	Me	H	12.6 ± 2.1
	73	Bn	Me	Me	Me	3.83 ± 0.7
	74	*i*Bu	CH_2(2-Naph)	*i*Bu	H	0.182 ± 0.02
	75	*i*Bu	CH_2(2-Naph)	H	H	82.0 ± 8.6
	76	*i*Bu	Bn	*i*Bu	H	3.5 ± 1.0
	77	Bn	Bn	*i*Bu	H	0.978 ± 0.171
	78	Bn	Me	Bn	Me	2.17 ± 0.14
	79	*i*Bu	Me	*i*Bu	Me	2.97 ± 0.15

1.6 Small Molecule α-Helix Mimetics

The α-helix is the most common secondary structure found within proteins, defined by a tight right-handed helical coil consisting of 3.6 residues per turn resulting in a pitch of 5.4 Å.[204] The side chains of the amino acids point outwards away from the helix in a staggered arrangement, minimising steric interference. The prevalence and regularity of α-helices in protein structure has made them interesting targets in the design of protein mimetics. The field of α-helix mimetics is especially important in the inhibition of PPIs, with over 62% of all PPIs featuring an α-helix at the binding surface.[205,206]

The primary aim in the design of α-helix mimetics is to design a scaffold that projects side chains at analogous distances and angular relationships to the side chains of residues in an α-helix. The scaffold needs to be conformationally controlled to ensure that it remains in the correct conformation for α-helical mimicry. Early examples of α-helix mimetics were focussed on the design of short conformationally restricted peptide sequences, with strategies such as hydrocarbon stapling, disulfide bridges, cross linking between residue side chains and hydrogen bonding routinely used in an attempt to control the conformation of unstructured peptides.[207] However, as discussed previously, peptide sequences have some major drawback, namely their metabolic instability and thus, attention was turned to identify small molecule α-helix mimetics. This section will discuss some of the success that has been reported in the design of small molecule α-helix mimetics.

1.6.1 Single Faced Small Molecule α-Helix Mimetics

The α-helix contains multiple binding interfaces, all of which can potentially form contacts with other proteins when engaged in a PPI. However, it has been found that the hot-spot residues are located on one recognition face of the α-helix for over 60% of all PPIs that feature an α-helix at the binding interface.[205] The three positions of the α-helix that were found by Arora *et. al.* to contribute most significantly to binding were the *i*, *i* + 4 and *i* + 7 positions (**Figure 33**). Arora *et. al.* used alanine scanning mutagenesis, which involves the sequential mutation of one amino acid in the sequence to alanine and the effect on binding affinity to its protein partner is monitored.[208] Any amino acid that when mutated to alanine results in a drop of binding affinity by more than 8.4 kJ mol^{-1} is deduced to be a hot-spot residue. Arora *et. al.* performed this analysis on 480 strongly interacting interfaces found within the PDB, enabling the identification of the most favourable hots-spot positions.[205]

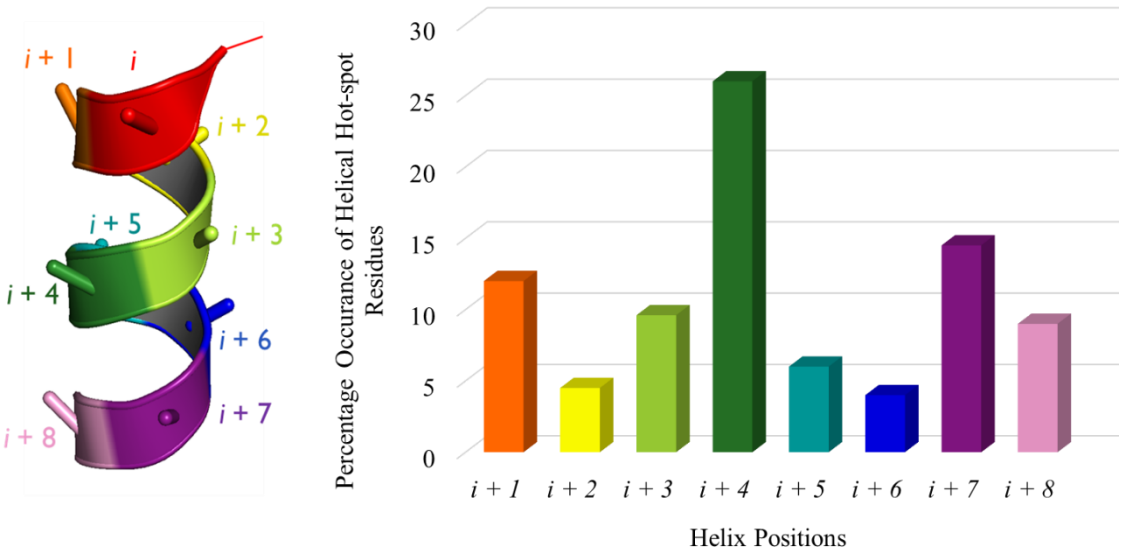


Figure 33. The α-helix and the frequency of positions occurring as a hot-spot residue. Figure reproduced from reference[205]

The investigation also revealed that hydrophobic residues were the most common residues found at binding interfaces with 34% of hot-spot residues being hydrophobic in nature. This was closely followed by residues with aromatic side chains, with 27% of hot-spots being aromatic. Polar or charged residues were also common, with 13% of amino acids with negatively charges side chains and 20% with positively charged side chains.[205] The finding reported by Arora *et. al.* supports the theory that PPIs are typically hydrophobic in nature.

The prevalence of PPIs that exist over one binding interface of an α-helix has led to the discovery of several small molecule α-helix mimetics that can mimic the *i*, *i* + 3/4 and *i* + 7 positions of the α-helix. The earliest example was reported by Hamilton *et. al.* in 2001, who designed terphenyl derivatives as structural and functional derivatives of the α-helix.[209] Hamilton *et. al.* have pioneered the field of small

molecule α-helix mimetics and many of the scaffolds that we see today bear resemblance to the original terphenyl scaffolds.

The terphenyl scaffold **80**, consists of three consecutive *ortho*-substituted phenyl groups. The three phenyl groups arrange themselves in a staggered conformation to reduce steric interactions between the *ortho*-substituents.[210] Although free rotation is still possible around the aryl-aryl bonds and other low energy conformations will be present in solution, it is believed that the staggered conformation is favoured due to the reduced steric clash between the *ortho*-substituents. Indeed, X-ray crystallography of a terphenyl derivative revealed a staggered conformation with dihedral angles of 59.1° and 120.7°.[209] The *ortho*-substituents project from the terphenyl core with similar angular relationships and distances to those of the *i*, *i* + 3/4 and *i* + 7 β-carbons in an α-helix (**Figure 34**).[211]

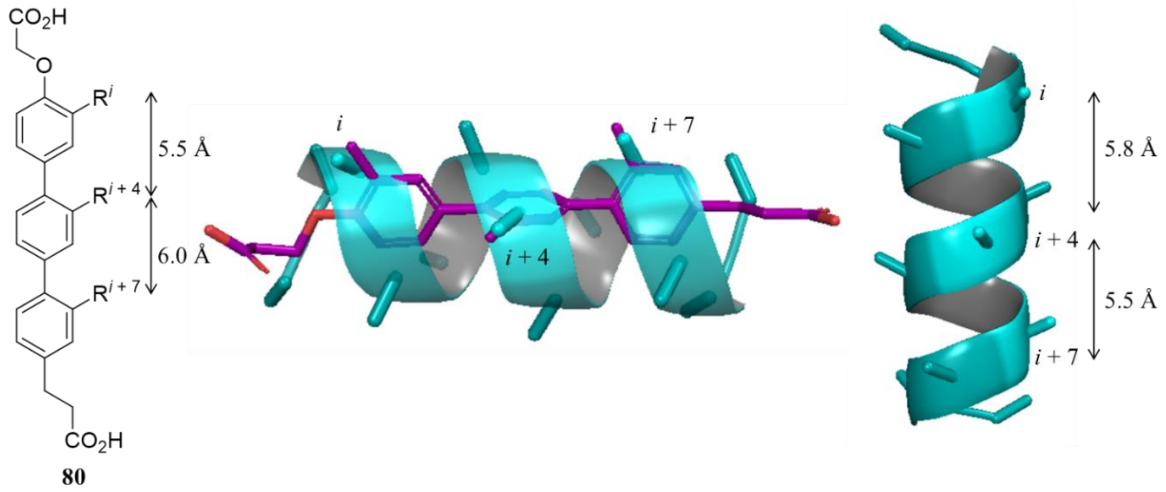


Figure 34. The terphenyl scaffold reported by Hamilton *et. al.* Overlaying the terphenyl scaffold with the α-helix, clearly shows excellent mimicry of the *ortho*-substituents with the *i*, *i* + 3/4 and *i* + 7 positions of the α-helix.

One of the major advantages of the terphenyl derivatives is that they can be synthesised in a modular fashion using sequential Negishi or Suzuki cross coupling reactions.[212] The reactions are high yielding and allow for the facile synthesis of libraries of terphenyls featuring different *ortho*-substituents by the simple substitution of the required coupling partner into the synthetic sequence. The terphenyl scaffold has been appended with suitable groups to inhibit a number of PPIs including the p53-Mdm2 PPI[203], the Bcl-xL-Bak PPI[213] and the interaction between CaM and smMLCK.[209] Although the terphenyl scaffold shows promise, they demonstrate very poor aqueous solubility and the low barrier of rotation between the aryl-aryl group renders them conformationally too flexible. Therefore, the original terphenyl scaffold has been subsequently modified, by Hamilton *et. al.* and many others.

The first modification to the terphenyl scaffold was introduced by Hamilton *et. al.* in an attempt to improve their solubility. By replacing the *ortho*-substituted phenyl groups with *ortho*-substituted pyridine groups the log*P* (o/w) dropped from 7.3 to 3.4 for the trimethyl substituted derivatives.[214] The

terpyridine derivatives **81** are expected to take on the same staggered conformation as the terphenyl derivatives, with the *ortho*-substituents mimicking the *i*, *i* + 3/4 and *i* + 7 positions of an α-helix. The biological activity of the terpyridines has however, not been explored by Hamilton *et. al.*

Hamilton *et. al.* are responsible for several classes of small molecule α-helix mimetics that can mimic one face of the α-helix. Including the terphenyls and the terpyridines, Hamilton *et. al.* have also reported the terephthalamides **82**[215], the tris-benzamides **83**[216], the diphenylindane **84**[217], the 5-6-5 imidazole-phenyl-thiazoles **85**[218], the biphenyl 4,4'-dicarboxamides **86**[219], the enaminones **87**[220] and the benzoylurea oligomers **88**[221]. All of these scaffolds project side chains that mimic the *i*, *i* + 3/4, *i* + 7 and in some situations the *i* + 11 positions of an α-helix (**Figure 35**). Heteroatoms are routinely incorporated to improve the aqueous solubility compared to the terphenyl scaffold. Additionally, hydrogen-bonding networks are often employed to ensure that the conformation that mimics the α-helix is the favoured conformation.[232,233]

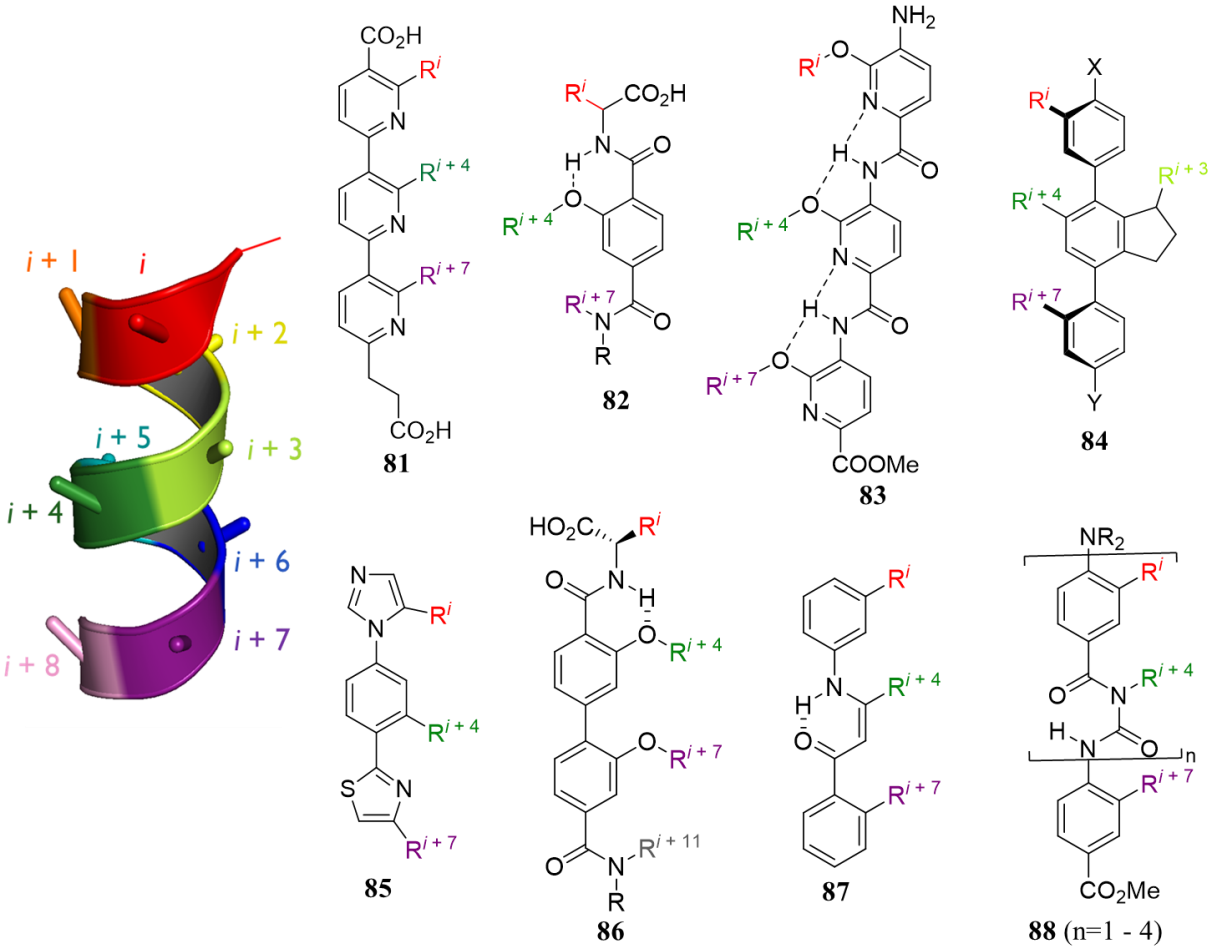


Figure 35. A collected of single faced α-helix mimetics designed and synthesised by Hamilton *et. al.*

Hamilton *et. al.* are not the only research group to explore the development of single faced small molecule α-helix mimetics. In 2019, Lim and co-workers, reported a bridged small molecule α-helix

mimetic, based on a triazine–piperazine–triazine scaffold.[224] The scaffold contains three functional side chains that match the distances and angular relationships between the *i*, *i* + 3/4 and *i* + 7 positions of the α-helix. Whilst the scaffold still bears a resemblance to those reported by Hamilton *et. al.*, a covalent bridge was introduced to reduce rotation around the central axis of the scaffold. The covalent bridge restricts the conformational freedom of the scaffold, addressing one of the major limitations of the terphenyls and other scaffolds reported by Hamilton *et. al.* A PEG bridge was introduced across the scaffold and the binding of the bridged analogue (**90**) and its parent compound (**89**)[225] to Mcl-1 was tested using a competitive fluorescent polarisation (FP) assay. Lim and co-workers were able to demonstrate that the covalently bridged analogous bind with a greater affinity and specificity to Mcl-1 than the non-bridged derivative (**Figure 36**). Lim *et. al.* propose that the increase in binding affinity is due to the entropic benefit of conformational restriction.[224]

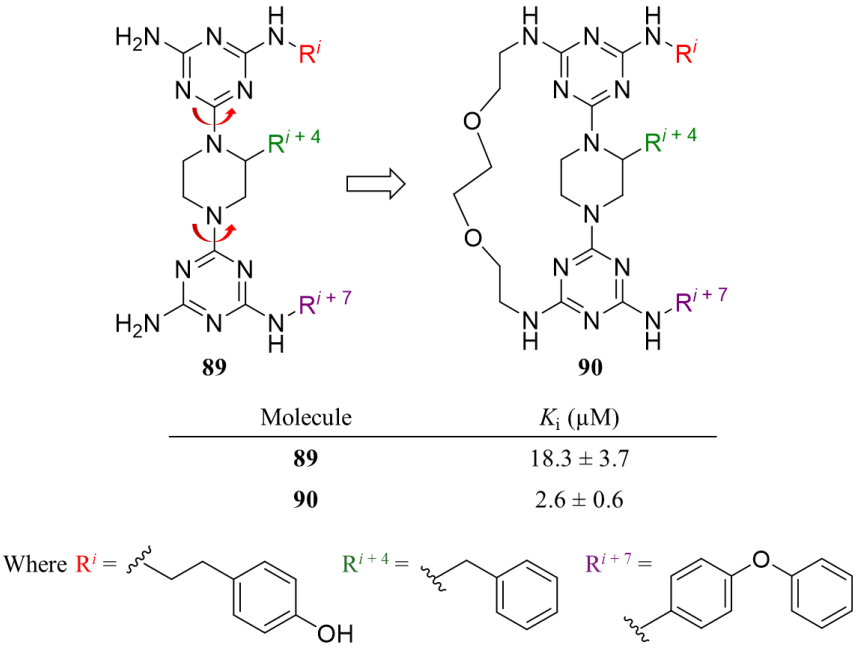


Figure 36. The triazine–piperazine–triazine scaffold reported by Lim *et. al.* The improvement in the binding affinity to Mcl1 following the incorporation of a PEG bridge is shown.

Many of the α-helical scaffolds that have been reported in the literature were inspired by the original terphenyl scaffold reported by Hamilton *et. al.* However, one example which was not inspired by the terphenyls is the polycyclic ether skeleton, reported by Hirama and co-workers.[226] Hirama *et. al.* took inspiration from the topological similarity between *trans*-fused polycyclic ether marine toxins and α-helices to design a ladder-like polycyclic ether scaffold as a new class of α-helix mimetics. Hirama *et. al.* noticed that the distance between two equatorial hydroxy groups on a 6/6/6/6/6 *trans*-fused pentacyclic ether scaffold (**91**) (4.8 Å) closely matches the distance between the side chains of the *i* and the *i* + 4 residues (~5 Å) (**Figure 37**). These hydroxy groups can be easily functionalised to incorporate desired side chains to mimic hot-spot residues. An advantage of the polycyclic ether scaffold is the

conformational rigidity that the scaffold will present due to the fused ring systems. Furthermore, the scaffold will likely display improved aqueous solubility than the terphenyls due to the introduction of oxygen atoms into the scaffold.

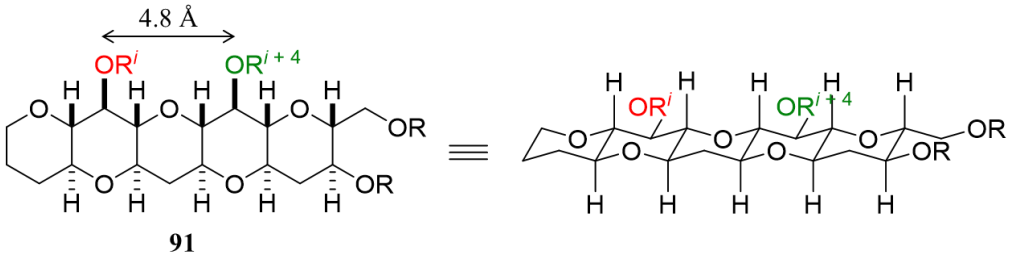


Figure 37. The polycyclic ether skeleton, reported by Hirama and co-workers. The distance between two equatorial hydroxy side chains, such as those highlighted, closely matches the distance between the *i* and *i* + 4 residues of the α-helix.

1.6.2 Multi-Facial Small Molecule α-Helix Mimetics

With over 60% of PPIs that feature an α-helix at the binding occurring over one face of the helix it is understandable that much of the focus on small molecule α-helix mimetics has been on developing scaffolds that can mimic the *i*, *i* + 3/4, *i* + 7 positions of the α-helix. However, roughly ~30% of PPIs utilize two faces of the α-helix for complexation while ~10% exist across all three faces of the α-helix (**Figure 38**). Therefore, there has also been a recent interest in the design of multi-facial small molecule α-helix mimetics. Historically, the design of multi-faced α-helix mimetics was limited to stabilised peptide helices as it is challenging to design a small molecule with two or more distinct binding faces.

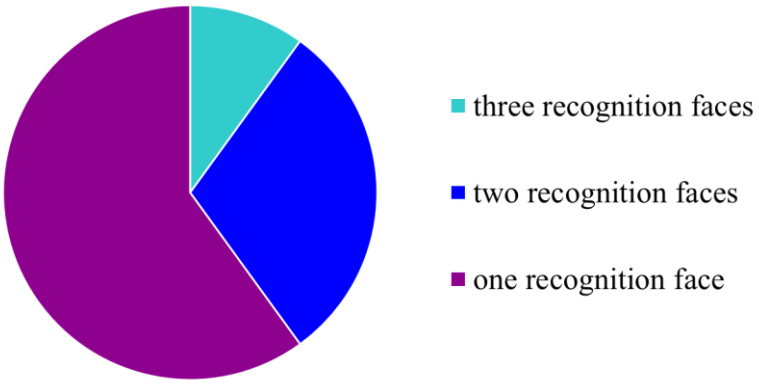


Figure 38. A pie-chart to show the percentage of PPIs that have hot-spot groups located on one, two and three recognition faces. Reference reproduced from reference[205]

Once again, Hamilton *et. al.* have pioneered the field of multi-faced α-helix mimicry, with some of the scaffolds that were originally designed to mimic one face of the α-helix derivatised to mimic multiple faces. For example, the tris-benzamide scaffold (**83**) reported by Hamilton *et. al.* in 2003[216] was redesigned to mimic the *i*, *i* + 2, *i* + 4 and *i* + 6 positions of the α-helix (**92**) (**Figure 39**).[227] They

proposed that intramolecular hydrogen bonding is responsible for the constrained conformation which places the side chains in the correct spatial orientation to mimic the desired positions. The synthesis of a bis-benzamide derivative was performed, and an x-ray crystal structure of a fully methylated derivative (R^{1-4} = Me) confirmed the intramolecular hydrogen bonding. Superposition of the x-ray structure and a representative α-helix yielded an RMSD of 0.43 Å, confirming excellent mimicry of the *i*, *i* + 2, *i* + 4 and *i* + 6 positions of an α-helix. Hamilton *et. al.* have also synthesised a hexa-substituted tris-benzamide (**93**), capable of mimicking the *i*, *i* + 2, *i* + 4, *i* + 6, *i* + 8 and *i* + 10 positions of an α-helix. This impressive result is possible due to the modular and scalable synthesis designed by Hamilton *et.al.*, which is amenable to extension, with much higher oligoamides able to be synthesised in theory.

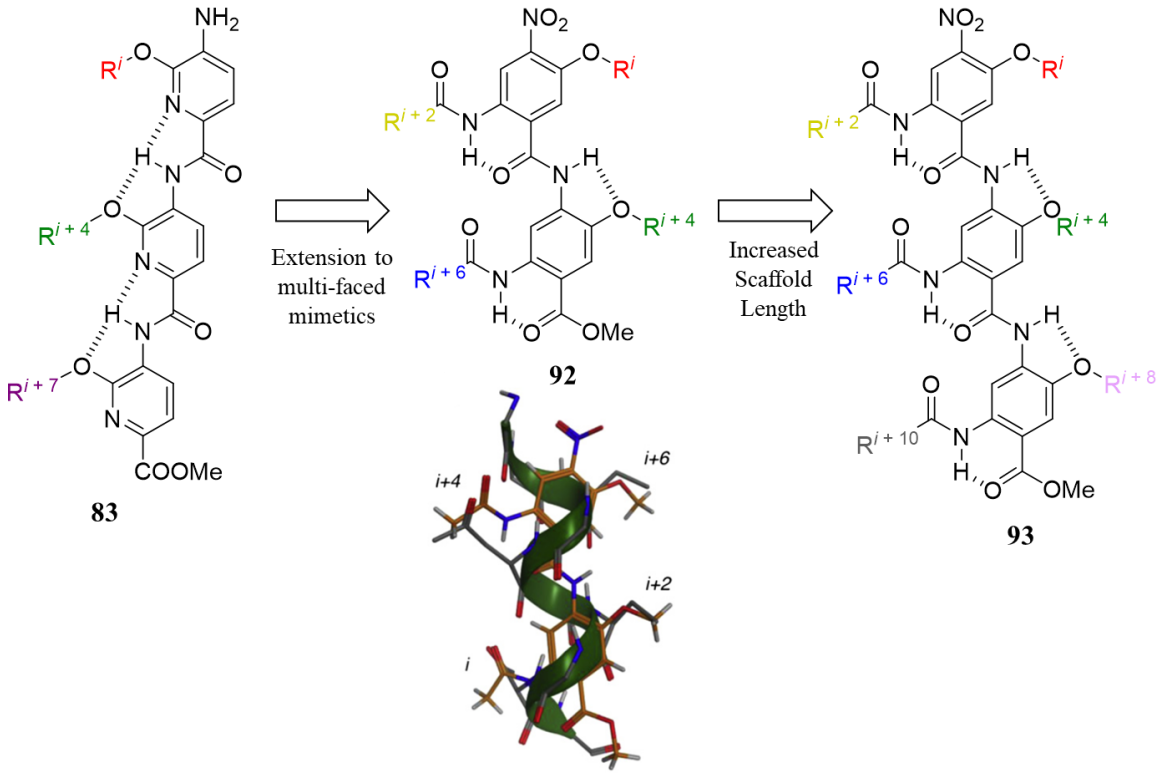


Figure 39. The evolution of the original tris-benzamide scaffold that can mimic one face of the α-helix, to the oligoamides that can mimic multiple faces of the α-helix. This figure was partly reproduced from reference [227]

A second scaffold originally reported by Hamilton *et. al.* as a single faced α-helix mimetic and since functionalised to mimic multiple faces of the α-helix is the benzoylurea scaffold (**87**). The original benzoylurea scaffold reported by Hamilton *et. al.*[221], features a conformationally controlled scaffold, held in a staggered conformation by intramolecular hydrogen bonding between the urea NH groups and the carbonyl oxygen atoms. Hamilton *et. al.* demonstrated that a benzoylurea scaffold could be used to design a mimetic of the *i*, *i* + 1, *i* + 4, *i* + 6 and *i* + 8 positions of an α-helix (**94**) (**Figure 40**).[228] The benzoylurea scaffold was synthesised in a modular fashion using dihydroxylated aromatics, which can be easily functionalised using alkylation reactions to incorporate both lipophilic and hydrophilic side

chains. An x-ray crystal structure of a representative benzoylurea confirmed the presence of intramolecular hydrogen bonding, responsible for holding the scaffold in the correct spatial orientation for α-helix mimicry. Superposition of the x-ray crystal structure with a model α-helix revealed excellent overlap of the substituents with the *i*, *i* + 1, *i* + 4, *i* + 6 and *i* + 8 residue side chains (RMSD = 1.25 Å).

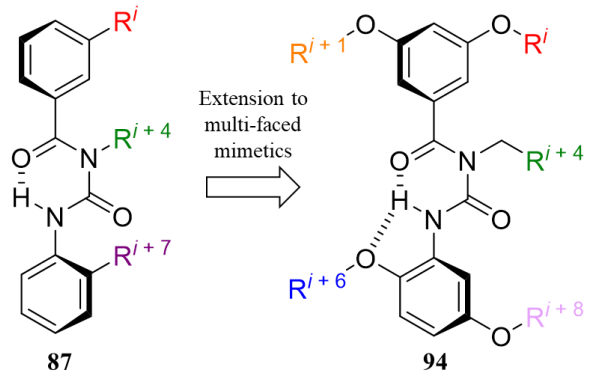


Figure 40. Modifying the original benzoylurea scaffold reported by Hamilton *et. al.* to mimic multiple faces of the α-helix.

Not all multi-faced α-helix mimetics were derived from single-faced α-helix mimetics. Instead Fletcher and co-workers drew inspiration from 1,2-diphenylacetylenes, common scaffolds used in the mimicry of β-turns and β-sheets.[229] Seminal work by Kemp *et. al.* demonstrated that 2,2'-functionalised diphenylacetylenes could mimic the β-turn by uni-directional β-sheet formation, held together by intramolecular hydrogen bonding.[230] The Fletcher group however, explored whether a similar diphenylacetylene scaffold could be used to induce α-helix mimicry (**Figure 41**). Four R groups are present which are intended to mimic the *i*, *i* + 2, *i* + 5 and *i* + 7 residues of an α-helix. Intramolecular hydrogen bonding was proposed to be key in maintaining the structural integrity required for α-helix mimicry. Substitution of the R groups with Me groups and superposition of a conformer of **95**, generated from an MD simulation, with an idealised polyalanine α-helix revealed an RMSD of 0.52 Å. This suggests that **95** mimics the *i*, *i* + 2, *i* + 5 and *i* + 7 residues of an α-helix to a high degree.

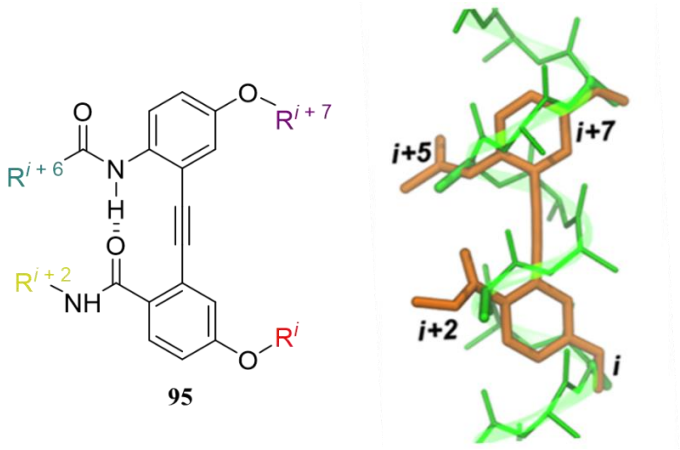


Figure 41. The diphenyl acetylene scaffold designed by Fletcher *et. al.* This figure was partly reproduced from reference [229]

A number of small molecule α-helix mimetic scaffolds have been reported in the literature, many of which were either invented by or inspired by the seminal work of Hamilton *et. al.* The majority of the scaffolds that current exist are only able to mimic one face of the α-helix, however, a transition towards multifaced α-helix mimetics, with more flexibility over the positions that can be mimicked, is present in the literature. Despite the progress that has been made in the field, a number of limitations are still yet to be overcome. The first is that often the synthesis of these scaffolds involves lengthy and low yielding routes. A second limitation of many of the scaffolds is the poor conformational control that they exhibit. Finally, the scaffolds that have been reported all mimic a set of residues very specific to that scaffold and little in the way of flexibility, in terms of the positions mimicked, is offered. A more desirable strategy would be to have one conformationally controlled scaffold, with multiple faces that can mimic any position desired by the incorporation of the appropriate side chain at the appropriate place on the scaffold. Such a scaffold could be described as a 'universal' α-helix mimetic.

1.7 Summary

The conformational preorganisation of therapeutics leads to higher binding affinities due to both entropic and enthalpic factors. This is especially important when targeting PPIs, which typically present featureless and large binding sites poorly amenable to the design of specific small molecule inhibitors. However, the consideration of free ligand conformation is often overlooked in SBDD. This is likely due to the challenges associated with measuring free ligand conformations of flexible molecules in solution. It requires a knowledge of the stereo and electronic factors that govern molecular conformation and the stereospecific synthetic methodologies available to synthesise conformationally controlled molecules. Finally, an understanding of the methods available to measure conformational behaviour in solution, including computational methods and NMR spectroscopy, is required. The multidisciplinary nature of this endeavour is perhaps the reason why it is often overlooked. This thesis presents the design, synthesis and conformational analysis of a new class of conformationally controlled α-helix mimetics using a combination of computational chemistry, organic synthesis and NMR spectroscopy. The multidisciplinary strategy is used to design a conformationally controlled mimetic of p53, and its binding to Mdm2 is also explored.

Chapter 2: Project Outline

The design of small molecule α-helix mimetics has been driven by the seminal work from Hamilton *et. al.*, who have reported the terphenyls as structural mimetics of the *i*, *i* + 3/4 and *i* + 7 residues of the α-helix.[209] Many of the scaffolds that have since been reported bear a resemblance to the original terphenyl scaffold. This thesis presents an entirely new class of small molecule α-helix mimetics devoid of strong similarity to the terphenyl scaffold. This thesis shows that conformationally controlled hydrocarbons, controlled through the avoidance of destabilising *syn*-pentane interactions, show promise as structural mimetics of the *i*, *i* + 3/4 and *i* + 7 residues of the α-helix.

In chapter 1, the design, synthesis and conformational analysis of molecules with tailored shapes using the iterative lithiation–borylation of boronic esters was discussed.[27] It was shown that an alternating *syn*-*anti* relationship between the methyl groups (**11**) yielded a linear conformational bias whereas an all-*syn* relationship (**10**) yielded a helical conformational bias, controlled by the avoidance of destabilising *syn*-pentane interactions. An investigation into the conformation of both **10** and **11** in solution using NMR spectroscopy and computation, revealed that both the linear (**11**) and the helical (**10**) scaffold present distinct faces from which methyl groups project. Furthermore, the distances between the methyl groups on one face of either **11** or **10** closely match the distances typically observed between the side chains of residues located at the *i*, *i* + 3/4 and *i* + 7 positions of the α-helix (5-7 Å) (**Figure 42**).

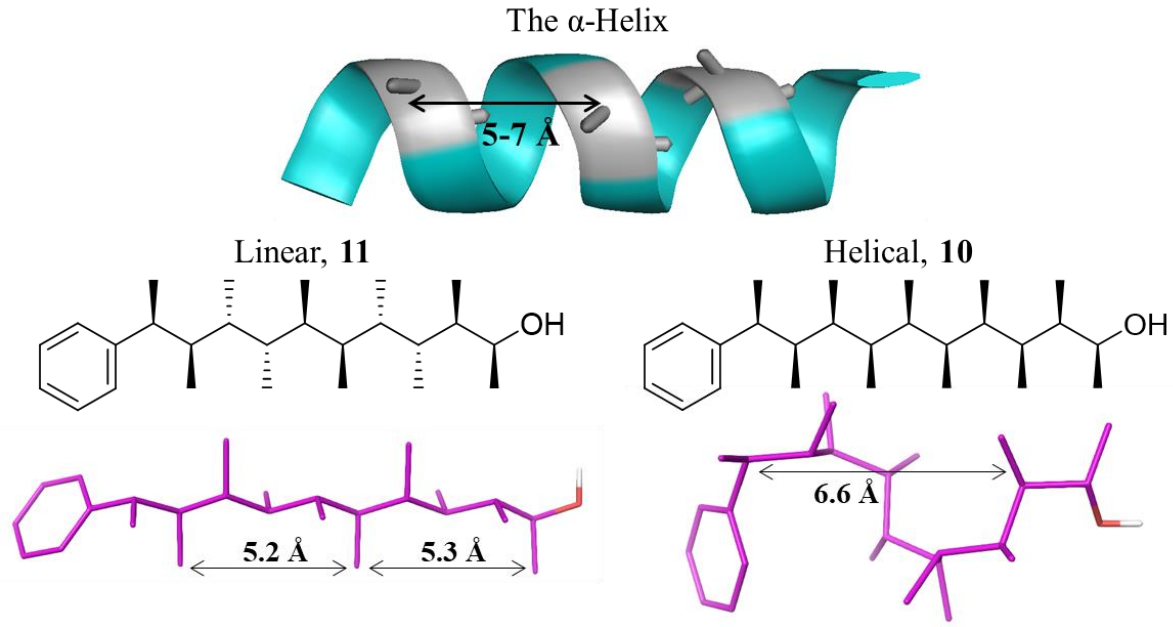


Figure 42. The distances between methyl groups on one face of both the linear and the helical scaffold are shown and compared to the distances between residues in an α-helix.

The access to a conformationally controlled scaffolds that can project methyl groups at analogous distances and angular relationships to those observed in the α-helix prompted the exploration of these scaffolds as a new class of small molecule α-helix mimetics. This forms the work of this thesis. Our focus in this PhD project has been on the alternating *syn-anti* diastereomer **11,** since the distances are a closer match to those typically observed in the α-helix. Additionally, significantly fewer homologations would be required to mimic the *i, i* + 3/4 and *i* + 7 residues for a linear scaffold compared to a helical scaffold. It is also challenging to have full control over the pitch of the helix for the all-*syn* diastereomer **10**. Superposition of the linear scaffold **11** with a representative polyalanine α-helix, clearly shows that the methyl groups found at the 1st, 5th and 9th position of the hydrocarbon scaffold closely match the methyl side chains of the *i, i* + 3/4 and *i* + 7 residues of the polyalanine α-helix (RMSD = 0.85 Å) (**Figure 43**). Therefore, the alternating *syn-anti* diastereomer **11** can be proposed to be suitable structural mimetic of the *i, i* + 3/4 and *i* + 7 residues of an α-helix.

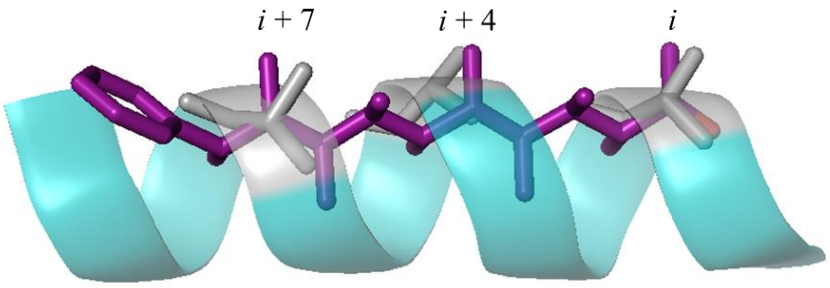


Figure 43. The superposition of the linear scaffold with a model polyalanine α -helix. The 1st, 5th and 9th positions of the linear scaffold **11** closely match the *i, i* + 3/4 and *i* + 7 residues of an α-helix.

Small molecule α-helix mimetics are an important class of targets in the inhibition of protein–protein interactions (PPIs). The hot-spot groups that dominate the binding between proteins engaged in a PPI are more often than not located on one face of an α-helical domain of one of the interacting proteins, and thus the methyl groups can potentially mimic these hot-spot groups. However, alanine is not typically found to be a hot-spot group as it cannot make extensive hydrophobic interactions with the binding partner. Therefore, groups other than methyl would need to be installed onto the scaffold to mimic a wider range of amino acids.

Fortunately, the iterative homologation of boronic esters, used to synthesise **11**, is easily amenable to the substitution of different groups on the scaffold in place of the methyl group. The linear scaffold **11** is synthesised using an assembly line process whereby an enantioenriched stannane undergoes a stereospecific tin-lithium exchange to yield a chiral carbenoid that can react with an electrophilic boronic ester to form a chiral boronate complex. Upon warming the boronate complex undergoes a stereospecific 1,2-metallate rearrangement to yield a homologated boronic ester that can be subjected to further homologations with no purification required in the intervening steps. The enantioenriched

stannane used in the synthesis of **11** resulted in methyl groups being incorporated alone the backbone. However, it is possible to incorporate different groups along the backbone using the same methodology by simply changing the side chains of the chiral carbenoid used in the sequence (**Figure 44**).

The enantioenriched stannanes used in the synthesis of **11** are generated from the stereospecific lithiation of ethyl triisopropylbenzoate (TIB) in the presence of *s*-BuLi and (±)-sparteine to form the chiral carbenoid. This is followed by electrophilic trapping with trimethyltin chloride to yield the α-stannyl benzoate in good enantiopurity.[27] The resulting stannane can be recrystalised in methanol to provide the stannane in >99.9:0.1 *e.r.* To change the group incorporated onto the scaffold all one must do is use a different benzoate ester in place of ethyl-TIB. If the resulting stannane is not a crystalline solid, and can thus not be recrystallised to high enantiopurities, then it is also possible to use the chiral ligand-assisted asymmetric deprotonation of the primary benzoate to generate the chiral carbenoid.[42] Benzoates can be synthesised by a simple Mitsunobu reaction of the corresponding alcohol, and thus a broad range of benzoates can be synthesised.[231]

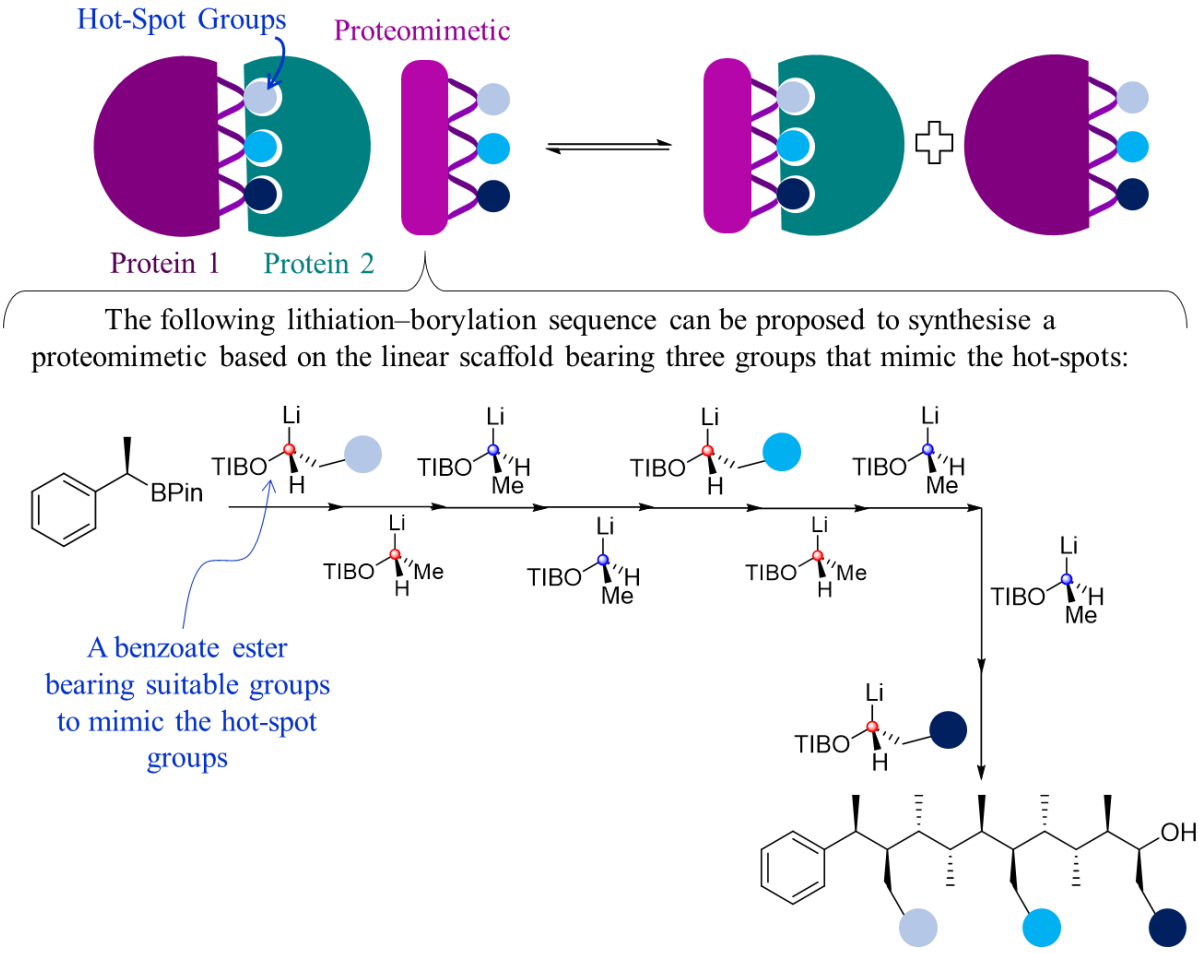


Figure 44. The proposed inhibition of a PPI using conformationally controlled hydrocarbons, synthesised using lithiation–borylation.

Therefore, using the iterative homologation of boronic esters it should be possible to synthesise conformationally controlled hydrocarbons bearing side chains at the 1st, 5th and 9th positions that mimic the side chains of the *i*, *i* + 3/4 and *i* + 7 residues of an α-helix. By the simple substitution of the benzoate ester, at judicious positions within the lithiation–borylation sequence, it is possible to install a hugely diverse set of side chains using one, high yielding and iterative methodology.

This thesis will first explore the effect on the conformational bias of the scaffold upon the introduction of extended side chains at the 1st, 5th and 9th position of the scaffold using computational methods including MM conformational searching. It is important to confirm that extending the methyl group by an additional methylene unit, does not disrupt the conformational bias by introducing additional *syn*-pentane interactions that disfavour the linear conformation. Once a suitable scaffold has been identified, the inhibition of the p53-Mdm2 PPI will be explored. p53 binds to a hydrophobic pocket on Mdm2, with the binding dominated by a Leu residue at *i*, a Trp residue at *i* + 4 and a Phe residue at *i* + *7*. This thesis will explore the design of a scaffold bearing appropriate groups to mimic the side chain of the Leu, Trp and Phe residues of p53 using MM conformational search calculations, such as scaffold **96**. The binding of designed p53 mimetics to Mdm2 will also be explored using protein-ligand docking calculations (**Figure 45**).

Following the design of potential p53 mimetics, the synthesis of the molecules will be performed by lithiation–borylation. Additionally, strategies to improve the aqueous solubility of the scaffold will be explored to allow for experimental validation of binding of the designed p53 mimetics to Mdm2.

Upon completion of the synthesis, a hybrid NMR spectroscopy and QM calculations workflow will be performed to examine the conformation of the designed p53 mimetic in solution. During this procedure, experimentally derived interproton distances, obtained from NOESY experiments, scalar coupling constants and chemical shifts are compared to Boltzmann averaged values calculated using QM calculations. The quality of the correlation between the two datasets will provide information regarding the population of conformers in solution.

Finally, the binding of the p53 mimetic to Mdm2 will be explored by ^{1}H-^{15}N HSQC spectroscopy. ^{15}N-labelled Mdm2 will be expressed and purified and the binding of the designed p53 mimetics will be confirmed by the presence of chemical shift perturbations (CSPs) in the ^{1}H-^{15}N HSQC spectra. The CSPs can be used to locate the binding of the ligand to Mdm2, confirming whether the ligand binds to the p53 pocket on Mdm2. Additionally, the CSPs can be tracked with increasing ligand concentration to generate binding curves, from which dissociation constants (K_d) can be estimated.

If this project is successful and demonstrates that conformationally controlled hydrocarbons can act as structural mimetics of p53 and bind to Mdm2, we hope that it will inspire the design of other α-helix mimetics, bearing different side chains to disrupt other therapeutically important PPIs.

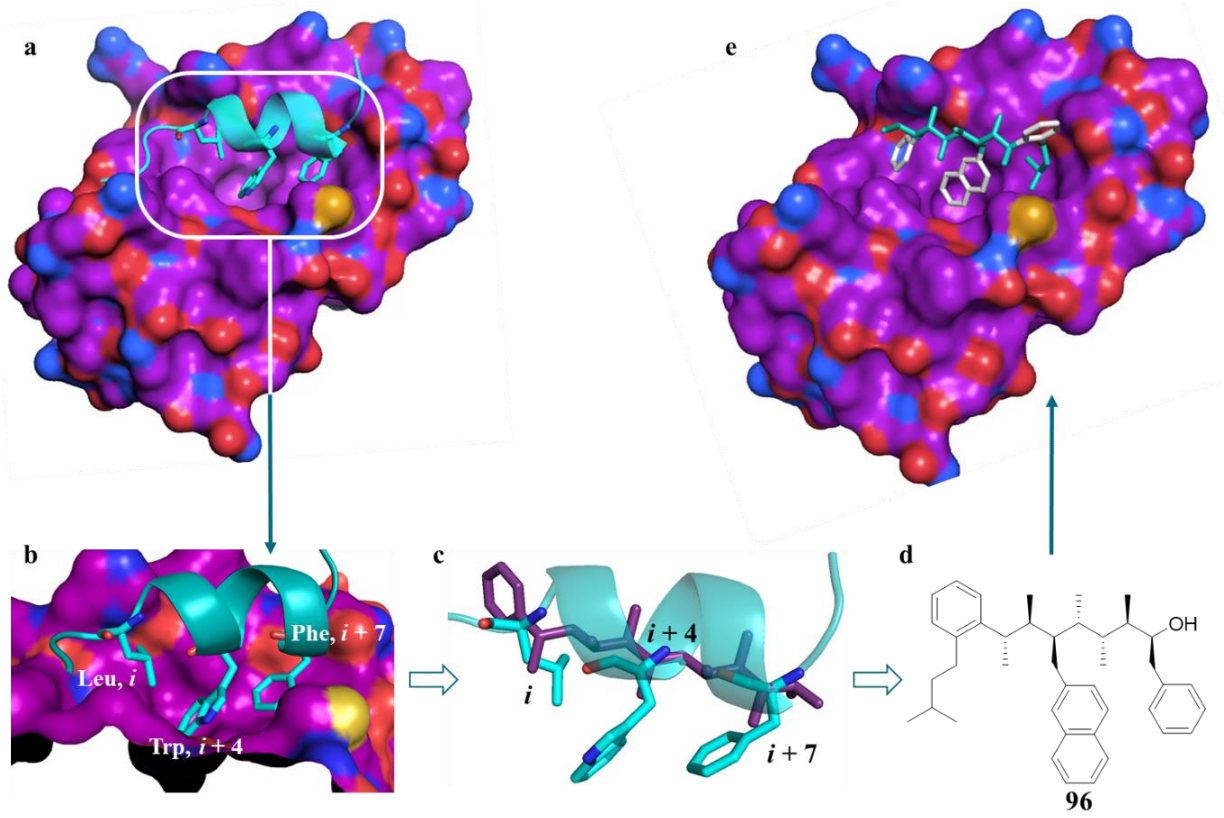


Figure 45. A summary of the proposed inhibition of the p53-Mdm2 PPI using conformationally controlled hydrocarbons **a**) The p53-Mdm2 PPI. The Mdm2 binding domain of p53 is shown in blue and Mdm2 is shown in purple. **b**) A zoom in of the p53-Mdm2 interaction, with the three hot-spot groups labelled. **c**) The overlay of the original linear scaffold **11** and p53. The 1st, 5th and 9th positions of the linear scaffold overlay well with the *i*, *i* + 4 and *i* + 7 residues of p53 as expected. **d**) A proposed p53 mimetic. The linear scaffold has been functionalised with suitable hot-spot groups. **e**) The proposed binding of a designed p53 mimetic, **96** to Mdm2. The binding of the designed p53 mimetic is similar to the binding of p53, however it is binding in the opposite direction as initially planned. PDB ID: 1YCR

Chapter 3: The Computational Design of an α-Helix Mimetic of p53

The principal aim of this project is to design a conformationally controlled scaffold that can mimic the key distances and angular relationships observed between residues located over one face of the α-helix. As discussed previously, many of the α-helix mimetics that have been previously reported suffer from poor conformational control, and a significant conformational reorganisation is required for binding, reducing the binding affinity. This conformational reorganisation can be avoided by the rational design of conformationally biased scaffolds, through the introduction of conformationally controlling groups onto the scaffold. This can include the introduction of *syn*-pentane interactions, capable of destabilising unwanted conformations and favouring the desired conformation, provided they are correctly introduced onto the scaffold.

Designing an alkane-based scaffold that is conformationally biased due to the avoidance of *syn*-pentane interactions can in theory be done using a simple diamond lattice analysis. This process involves manually drawing the alkyl chain onto a diamond lattice, and then redrawing conformers with certain bonds rotated to explore the potential *gauche* and *syn*-pentane interactions (**Figure 46**). Although this is a useful tool for quickly identifying conformations containing *syn*-pentane or *gauche* interactions, it does not provide quantitative information regarding the populations of various conformations in solution. It is also a very laborious task, requiring many manual drawings of conformers, and conformations can be easily missed in molecule with a large number of rotatable bonds.

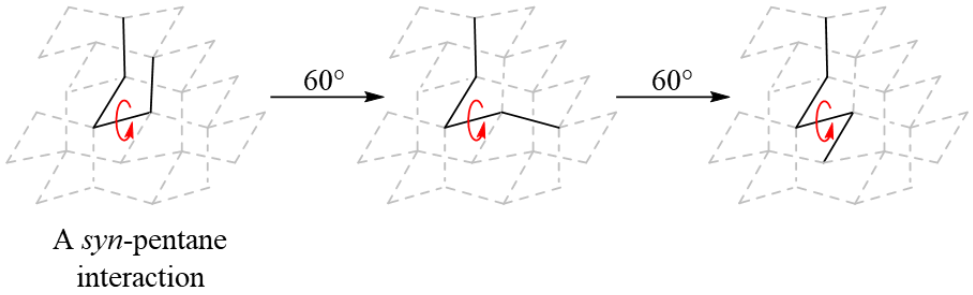


Figure 46. The diamond lattice analysis of 3 conformers of pentane. The first conformer shown features a *syn*-pentane interaction however using the diamond lattice it is clear that rotation of this final dihedral by 60° leads to a conformer with no *syn*-pentane interaction present. This process can be repeated. The diamond lattice can help to identify those conformers with no *syn*-pentane interactions.

An alternative approach is to use computational methods such as molecular mechanics (MM) conformational searching. This chapter will discuss the design of an α-helix mimetic of p53 using both the diamond lattice and MM conformational searching. Additionally, this chapter will explore the binding of the designed p53 mimetic to Mdm2 using molecular docking.

3.1 Using the Diamond Lattice to Investigate Molecular Conformation

As discussed in chapter 2, the alternating *syn-anti* diastereomer (**11**) of the contiguously methyl substituted scaffolds designed and synthesised by Aggarwal *et. al.* favours a linear conformation in solution.[27] The methyl groups at the 1st, 5th and 9th positions can mimic the residues at the *i*, *i* + 3/4 and *i* + 7 position of an α-helix and therefore the linear scaffold could act as a suitable α-helix mimetic. The conformation of the molecule is controlled by the avoidance of destabilising *syn*-pentane interactions between the methyl groups. Drawing the linear conformation of a *syn-anti* diastereomer of a poly-methylated hydrocarbon (**97**) onto a simple 2D-diamond lattice it can clearly be seen that no *syn*-pentane interactions exist and thus this conformation is favoured (**Figure 47**). If the molecule were to adopt any conformation other than linear, at least one *syn*-pentane interaction would be incurred, reducing the population of these conformers to <0.5% at room temperature.

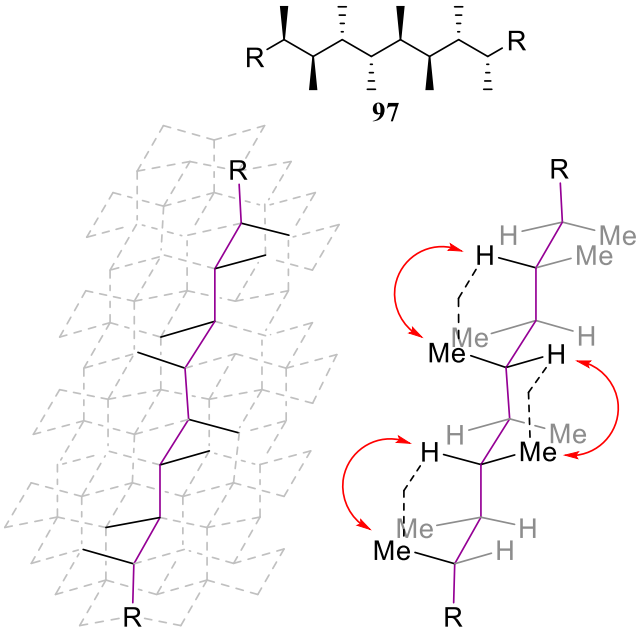


Figure 47. The diamond lattice analysis of a linear scaffold **97**. The conformational bias of the molecule is such that it avoids all syn-pentane interactions (red arrows).

Using the linear scaffold **97** as a starting point, the diamond lattice analysis was used to explore the effect on the conformation upon adding hot-spot groups at the 1st, 5th and 9th positions of the scaffold. If the chain length of a methyl group is increased by one carbon, resulting in structure **98**, it can clearly be seen using the diamond lattice, that there is no position for the additional carbon that avoids a *syn*-pentane interaction (**Figure 48b**). However, if an adjacent methyl group is removed from the scaffold then a linear conformation can exist with no *syn*-pentane interactions (**Figure 48c**). Additionally, if a third carbon is added to the side chain, such as structure **99**, there is only one position on the diamond lattice for this atom that avoids a *syn*-pentane interaction. Thus, by removing one adjacent methyl group the side chain is controlled by up to three bonds.

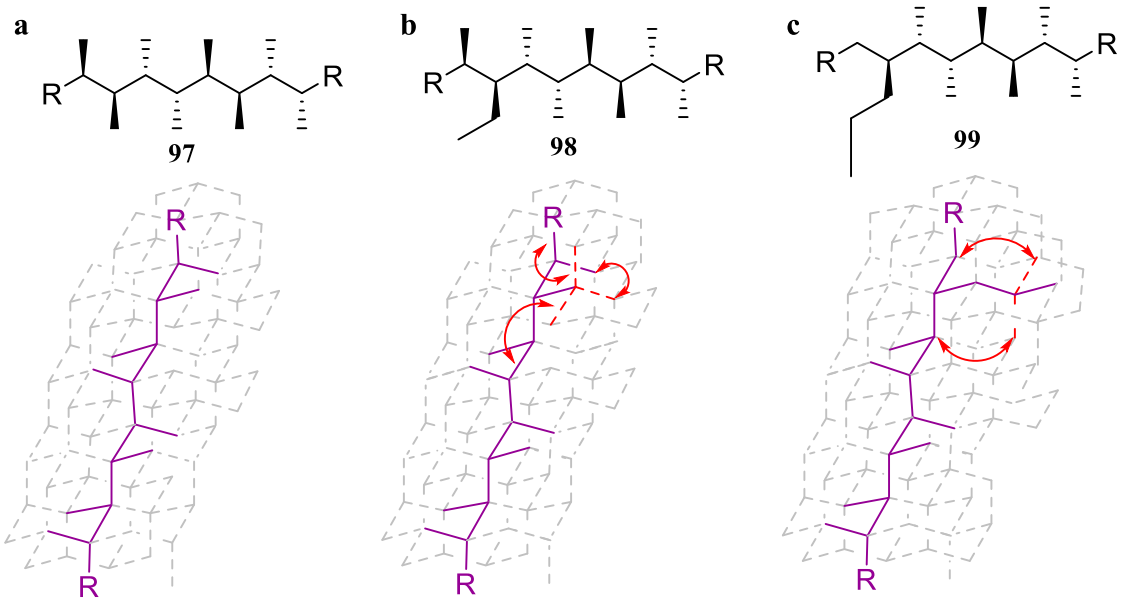


Figure 48. The diamond lattice design of an α-helix mimetic. **a)** The conformation of the linear scaffold **b)** Increasing the length of a methyl group by one carbon results in a syn-pentane interaction **c)** Removing an adjacent methyl group prevents a syn-pentane interaction.

Based on the diamond lattice analysis, a suitable scaffold that can mimic the *i*, *i + 3/4* and *i + 7* positions of the α-helix is shown in **Figure 49**. Methyl groups have been removed at adjacent positions to each mimetic side chain. The linear conformation of this molecule is predicted by the diamond lattice analysis to avoid any *syn*-pentane interactions, resulting in a linear conformational bias. In order to confirm this conformational bias and calculate the population of conformers adopting a linear conformation, MM conformational searching needs to be employed.

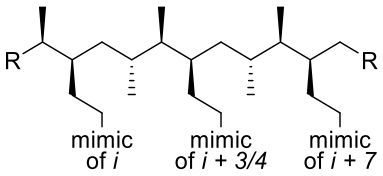


Figure 49. A proposed linear scaffold capable of mimicking the *i*, *i* + 3/4 and *i* + 7 positions of the α-helix.

3.2 Molecular Mechanics Conformational Searching to Explore the Conformational Space of a Designed p53 Mimetic

To investigate whether the conformational bias of the molecule is as predicted by the diamond lattice analysis, and to calculate the potential energies and corresponding conformer populations, a MM conformational search was performed. The conformational search was performed using MacroModel, a software package available from Schrödinger.[232] As discussed in section 1.3.2 of chapter 1, a conformational search locates a number of low energy conformers of a molecule and estimates their potential energies which can be translated to a Boltzmann population.

A Monte Carlo Multiple Minimum (MCMM) conformational search was used in this study. The MCMM conformational search method is a stochastic search method whereby conformers are generated by randomly rotating either the cartesian coordinated or the torsional bonds by differing amounts.[233] The distorted structure is minimised and compared to the structures generated in previous conformational search steps. The minimised conformer is either stored as a new structure or eliminated on the groups of its energy or redundancy with previous conformers found. The MCMM method was chosen based on its ability to explore both close and distance areas of the potential energy surface. This is very important when studying the conformational landscape of flexible molecules with a large number of rotatable bonds. The conformers generated from the MCMM conformational search were minimised using the Truncated Newton Conjugate Gradient (TNCG) minimisation method. The TNCG method uses the gradient of the energy function to identify convergence. The TNCG method is very efficient at diverting the direction of the minimisation away from nonconvex regions towards low gradient structures, reducing the number of imaginary frequencies found.[234,235] The energy of the conformers was estimated using the Merck Molecular Force Field (MMFFs) force field, which was developed to accurately treat non-bonding interactions.[236] Due to the potential for *syn*-pentane interactions, a non-bonding interaction, within the molecule contributing to their energy, the MMFFs force field was chosen. A full description of the MM conformational search parameters used in this study is provided in the supporting information.

An MCMM conformational search was performed on a representative structure (**100)** with propyl side chains representing the hot-spot groups. By simplifying the hot-spot groups to propyl side chains the computational time taken for the calculation is reduced but information about the conformational bias of the backbone is still obtained. The results of the conformational search are summarised as a bubble plot whereby each dihedral angle along the backbone of the molecule is represented by a bubble whose size is related to the population of the conformer corresponding to that dihedral angle (**Figure 50**). For a linear conformation, the dihedral angles along the backbone of the molecule should all be *anti* or ~180°. It is clear from the graph in **Figure 50** that this is not the case and structure **100** does not preferentially adopt a linear conformation. Instead, alternating dihedral angles of *gauche-anti-anti-gauche-gauche-anti-anti etc* are observed, resulting in a kinked conformation. This was unexpected as the linear conformation contains no *syn*-pentane interactions, however, the lowest energy conformer found by MM also has no *syn*-pentane interactions.

If **100** was controlled by the avoidance of *syn*-pentane interactions alone, a roughly equal weighting of the kinked conformer and the linear conformer would be expected, since they both contain no *syn*-pentane interactions so one should not be favoured over the other. Additionally, the two conformers contain the same number of *gauche* interactions. This effect can be explained by a deviation of the dihedral angles away from those assumed from a diamond lattice analysis. This distortion of the dihedral

angles from the ideal values, reduces the energy penalty of any destabilising interactions such as *syn*-pentane or *gauche* interactions. This makes it impossible to discriminate between conformers with the same number of destabilising interactions using a diamond lattice analysis.

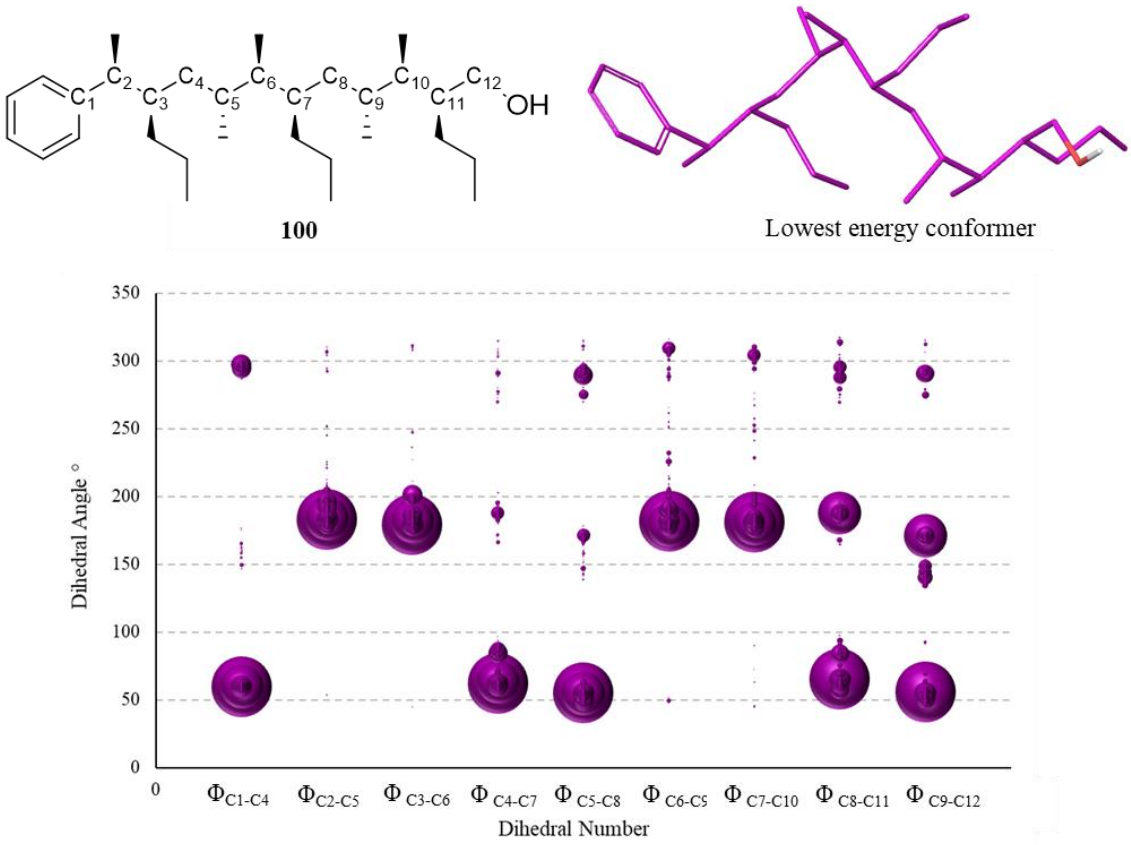


Figure 50. The bubble plot of representative structure **100**. The plot clearly shows that a linear conformation is not favoured.

This problem can be addressed by investigating what would happen to the conformational landscape if the adjacent methyl groups are not removed. It may then be possible to find ways of destabilising the preferred conformation and favouring the linear conformation. Interestingly, it was found that a linear conformation was preferred for the fully-methylated structure **101**, despite incurring multiple *syn*-pentane interactions. The bubble plot in **Figure 51** clearly shows a bias towards a linear conformation, with dihedral angles of ~180° along the backbone. Totalling the contributions from all conformers adopting a linear conformation gave a total of 73%.

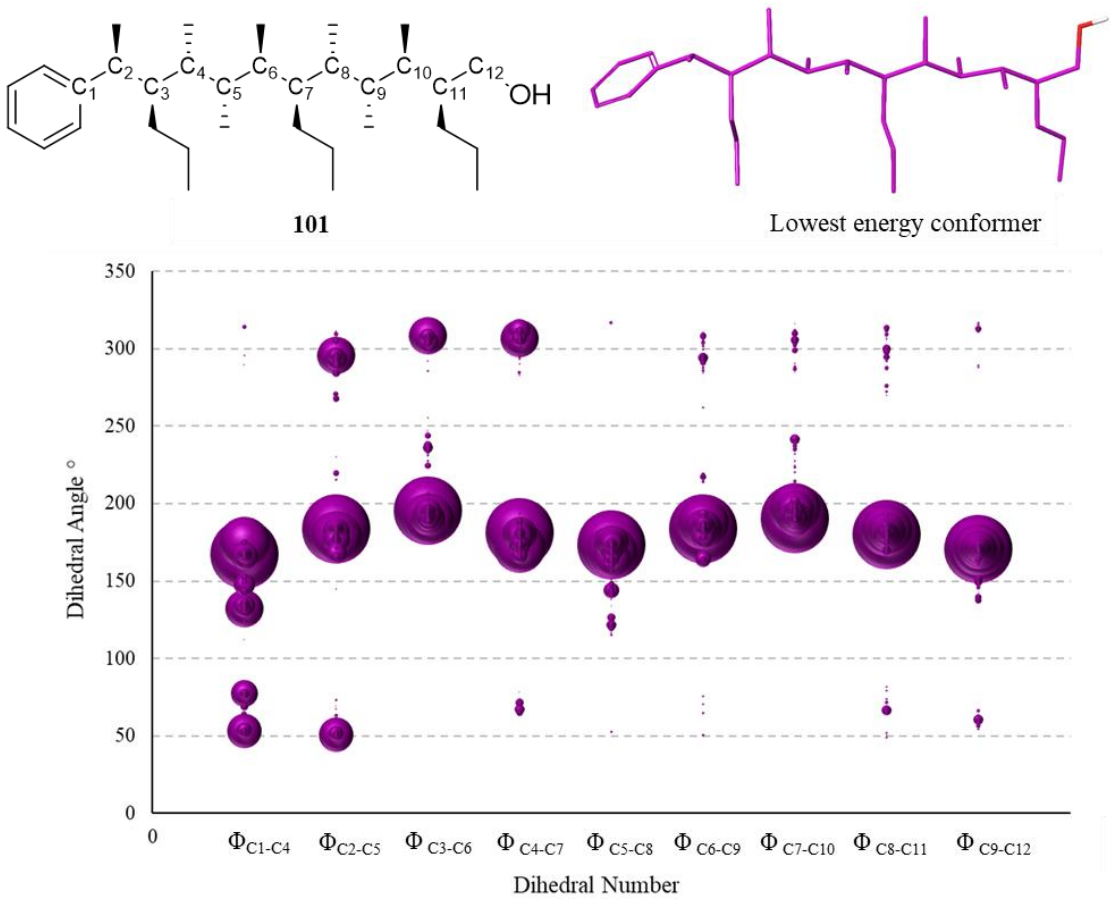


Figure 51. The bubble plot from a MM conformer search of **101**. Dihedral angles of ~180° are favoured, resulting in a linear conformational bias.

The linear bias that is observed for **101** is unexpected because of the *syn*-pentane interactions present. However, a closer inspection of the linear conformation of **101** reveals that, although two *syn*-pentane interactions are present, the dihedral angles within the molecule have distorted away from their ideal values. This offsets one of the *syn*-pentane interactions by ~40° (**Figure 52**). This will reduce the energy penalty of the *syn*-pentane interaction considerably making the linear conformation more favourable. Importantly, the other conformers that are populated contain two or more *syn*-pentane interactions, which do not exhibit such a considerable offset between the groups involved. This will raise the energy of the other conformations, resulting in the linear conformational bias observed.

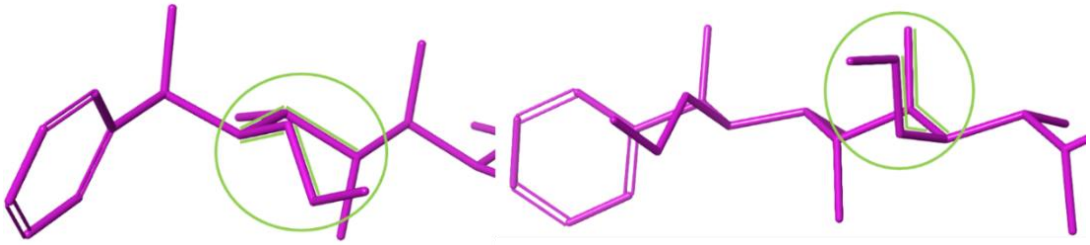


Figure 52. The two syn-pentane interactions present in the linear conformation of **101**. The one highlighted on the right is offset by ~40°.

Although structure **101** demonstrates the conformational bias desired, synthetically it is very demanding with ten lithiation–borylation reactions to be performed. The length of the mimetic can be shortened by incorporating one of the propyl side chains directly onto the aromatic ring as shown in structure **102** (**Figure 53**). By doing this, the increased conformational control of planarity of the aromatic ring over an sp^3-hybridised methine centre is also exploited. A MM conformational search of **102** revealed that a linear conformation is preferred with over 94% of the conformer population adopting a linear conformation. The increased bias towards a linear conformation, for structure **102** compared to structure **101**, could be attributed to the fact that the linear conformation of **102** only contains one *syn*-pentane interaction whilst other conformations have two or more. Additionally, shortening the chain reduces the entropic penalty incurred on conformational organisation.

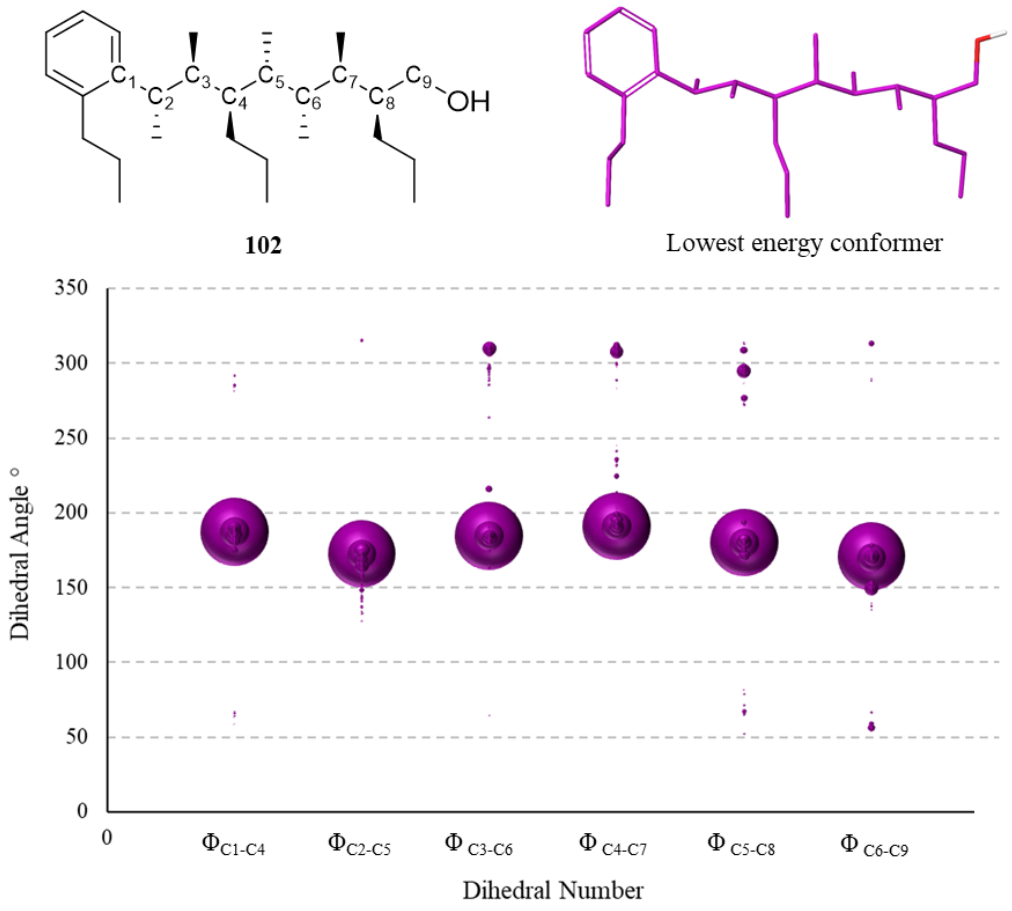


Figure 53. The bubble plot of structure **102** reveals a linear conformation is preferred with over 94% of the conformer population adopting a linear conformation.

To confirm whether the designed scaffold mimics the key distances and angular relationships to those observed in p53, the lowest energy conformation of structure **102** was overlaid with the Mdm2 binding domain of p53. This revealed that the propyl side chains closely align with the side chains of the hot-spot groups of p53. Using the pair fitting function in PyMol, and fitting the three carbon atoms on the backbone of **102** highlighted in **Figure 54** with the C_α of the hot-spot amino acids revealed an RMSD

of 0.60 Å. An RMSD of <1 Å is considered to be a strong match between two structures and confirms that our designed scaffold closely matches the *i*, *i* + 4 and *i* + 7 positions of p53.

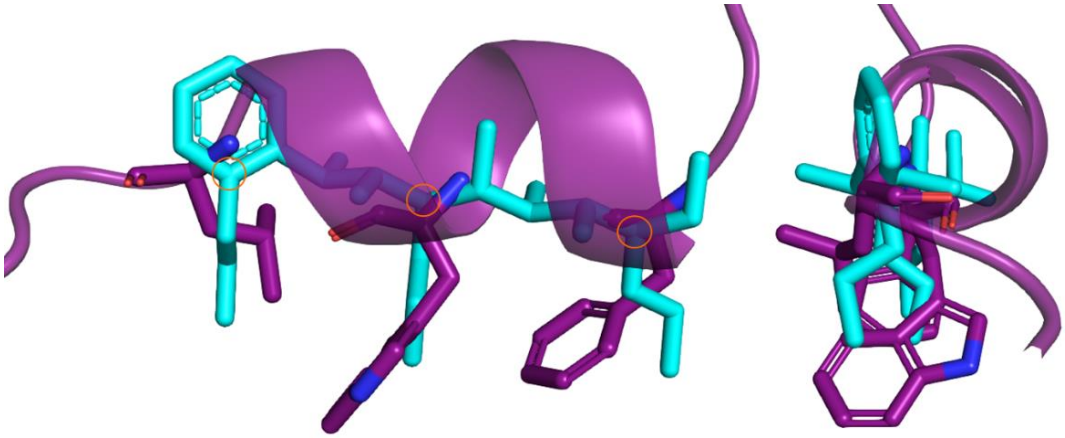


Figure 54. An overlay of the designed scaffold **102** with p53. The atoms used in the pair fitting are highlighted by orange circles.

With a suitable scaffold designed, the effect on the conformational bias upon the addition of suitable hot-spot groups was explored. A total of three hot-spot groups exist on p53; a leucine residue at *i*, a tryptophan residue at *i* + 4 and a phenylalanine residue at *i* + 7. The designed p53 mimetic **96** is shown in **Figure 55**. An isopropyl group has been incorporated to mimic the leucine side chain, a naphthyl group is in place to mimic the indole ring of the tryptophan side chain and a benzyl group is present to mimic the phenylalanine side chain. Although the indole NH of tryptophan in p53 is engaged in a hydrogen bond with Leu54 of Mdm2[179], it has been reported by Hamilton *et. al.* that a naphthyl group is a suitable mimetic of the indole of tryptophan.[203] Using a naphthyl group in place of an indole group avoids the need for NH protection and deprotection during the synthetic procedure, making it synthetically more attractive. The scaffold was further simplified by shortening the length of the scaffold through the removal of the terminal methylene group since adjacent methylene group.

The hot-spot groups are large and contain aromatic groups that could interact via π-π stacking. To ensure that these groups do not perturb the desired conformational bias of the scaffold, a MM conformational search of structure **96** was performed (**Figure 55**). The MM conformer search of structure **96** reveals a linear conformational bias is still observed, with over 74% of conformers adopting a linear conformation.

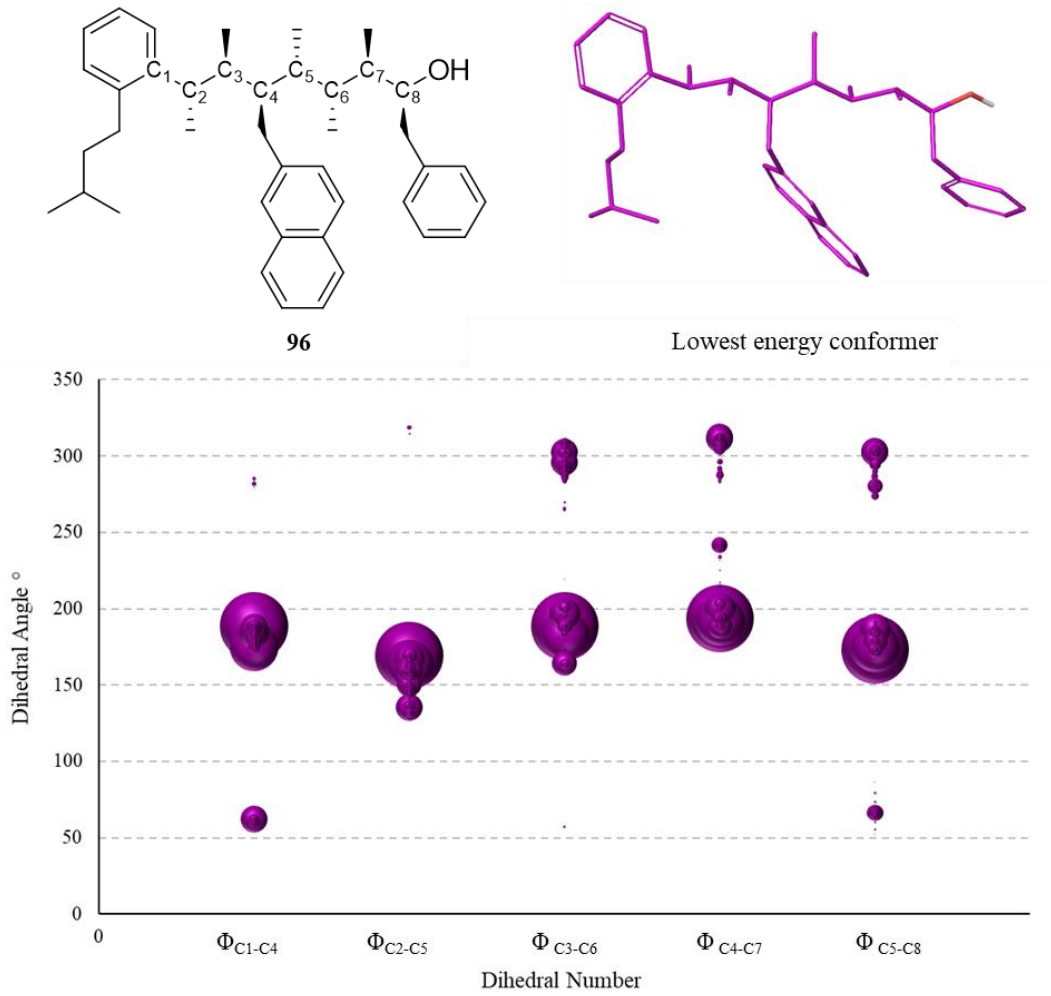


Figure 55. The bubble plot of the designed p53 mimetic **96**. The linear conformational bias is retained upon addition of the hot-spot groups with over 74% of the conformer population adopting a linear conformation.

The conformational control of the molecule only extends along the chain of the hot-spot groups by one carbon atom, meaning that the position of the side chains is flexible. For example, an overlay of the three lowest energy conformers, which have a combined population of ~64%, reveals a linear backbone with differing positions of the hot-spot groups **(Figure 56**). This side chain flexibility is advantageous for binding to Mdm2, as it will allow the hot-spot groups to adopt the position needed to maximise the binding affinity between the designed mimetic and Mdm2, with little increase in the free energy of the mimetic. Additionally, the conformational mobility of the side chains will improve the mobility of the mimetic across a cell membrane.[237]

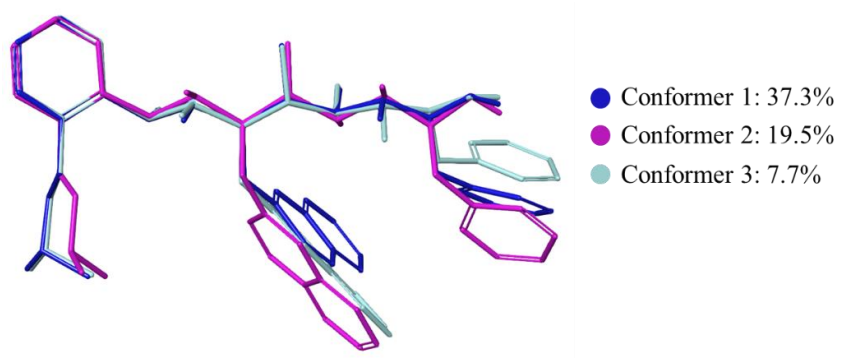


Figure 56. The three lowest energy conformers of the designed mimetic **96** showing the side chain flexibility.

We have now successfully designed a p53 mimetic (**96**) which places the hot-spot groups at key distances and angular relationships to mimic those of p53. Although a mimetic of p53 has been designed, the basic scaffold designed is capable of mimicking the *i*, *i* + 3/4 and *i* + 7 positions of an α-helix, and simple substitution of the hot-spot groups, using lithiation–borylation, provides access to a universal single faced α-helix mimetic.

3.3 Exploring Binding of the Designed p53 Mimetic 96 to Mdm2 using Molecular Docking

Following the successful design of a conformationally controlled p53 mimetic (**96**) the binding of the designed mimetic to Mdm2 was explored using molecular docking. Molecular docking, also known as protein-ligand docking, predicts how strongly a molecule binds to a receptor, typically a protein. In addition, it predicts the preferred pose of the ligand in the binding site of the receptor. Protein-ligand docking is routinely used in the drug discovery process, most typically it is used to screen large virtual libraries of typical drug molecules against a receptor. Once suitable candidates have been identified from the virtual screening, the structure is revised and functional groups altered to maximise the binding interaction between the ligand and the protein, predicted by protein-ligand docking. This process is known as structure-based drug design (SBDD).[238]

A number of challenges exist in the accurate prediction of protein-ligand docking and as a result a large number of docking programs are available that all aim to address these challenges in differing ways. There are two key challenges in molecular docking that docking programs routinely address and aim to improve. The first is the generation of ligand poses i.e. correctly predicting the possible conformations of the ligand, and the energy of that conformation, along with the fit between this pose and the binding site. It is not uncommon for small molecules to contain many conformational degrees of freedom and accurately incorporating this ligand flexibility into the calculation is not straightforward.[239] Ideally, a simulation would account for both the bound and the unbound conformation of the ligand and account for any conformational reorganisation that is needed to move from the unbound to the bound, however

molecular docking simulations are currently unable to do this. As well as limitations in ligand flexibility, molecular docking simulations do not typically consider receptor flexibility. However, it is known that proteins are conformationally labile and can select different conformations to maximise affinity with a small molecule.[10]

The second important parameter is the scoring function used, which predicts the energy of the interaction and ultimately which pose of the ligand is preferred. The aim of the scoring function is to provide accurate predictions of ligand binding in reasonable computational time.[240] Advancements in computing power has made this aim more realistic however, limitations within the scoring functions of many molecular docking programs exist including correctly accounting for entropy, solvation and receptor flexibility.[241] Despite the limitations associated with molecular docking, molecular docking programs are able to predict binding affinities with errors of roughly 8.5-13 kJ mol^{-1} provided the receptor protein does not exist in multiple conformations.[242]

Several molecular docking programs exist, such as Glide, GOLD and FlexAID. However one of the most successful and well cited docking software programs is AutoDock.[243] AutoDock is a preferred choice for many drug discoverers as it enables ample sampling of the conformational space of both the ligand and receptor. This is achieved by allowing selected side chains of the receptor to be flexible and not fixing the bond length or angles within the receptor, two common limitations of other docking software.[242]

AutoDock Vina, a development of the initial software with improved speed and accuracy, was chosen for performing the ligand-protein docking calculations. An additional benefit of AutoDock Vina is the scoring function used. Most scoring functions will be based on only one of the classes of scoring functions, of which the most typical are force field, empirical or knowledge-based. Vina however uses a hybrid of both the empirical and knowledge-based classes. Empirical based scoring functions combines the contributions of various types of interactions between two partners while knowledge-based scoring functions reproduce the rules and principles statistically derived from experimentally determined structures. The scoring function from AutoDock Vina uses empirical information from conformational preferences of receptor-ligand preferences and experimental binding affinities[244] The scoring function of AutoDock Vina is shown in **Equation 5**.

$$c = \sum_{i<j} f_{titj}\,(r_{ij})$$

Equation 5

Where f_{titj} is a symmetric set of functions and r_{ij} is the interatomic distance between atoms i and j. The summation is performed over all pairs of atoms that can move relative to one another. The scoring

function aims to find the global and local minima of *c* and rank them. AutoDock Vina was found to perform best for ligands with 8 or more rotational bonds.[245,246] Many groups have compared the performance of the most popular molecular docking programs (such as AutoDock Vina, GOLD, Glide and LeDock) and they have all found AutoDock Vina to perform well and more often than not delivers the most accurate binding affinities, attributed to the superior scoring function.[246–248]

To perform a docking simulation in AutoDock Vina the receptor and the ligand are prepared using AutoDock Tools, a graphical user interface for coordinate preparation, docking and analysis.[242] The lowest energy conformation of the designed p53 mimetic **96**, found from the MM conformational search, was chosen as the starting coordinates for the ligand and the receptor coordinates were taken from a crystal structure. The choice of crystal structure, from which the receptor coordinates are taken, is an important parameter to consider. A crystal structure of Mdm2 and the Mdm2 binding domain of p53 exists (PDB: 1YCR), however a number of crystal structures of Mdm2 with small molecule or peptide therapeutics also exist (PDB: 4HFZ, 1RV1, 4ERF). Examination of these crystal structures reveals that the conformation of the binding site on Mdm2 changes slightly depending on the ligand/peptide bound to it. For instance, when p53 is bound to Mdm2, a much more open binding cleft was observed compared to small molecule inhibitors of Mdm2. This is not unexpected, as it has been reported that the p53 binding site of Mdm2 is moderately conformationally labile.[249–252] Although AutoDock Vina allows for receptor side chain flexibility, it is not adequate enough to sample the conformational flexibility of the p53 binding site of Mdm2. Thus, it is important to use an appropriate crystal structure, with a p53 binding site conformation similar to that expected with the designed ligand. Since a mimetic of p53 is being designed, the crystal structure of p53 and Mdm2 (PDB: 1YCR) was chosen for the molecular docking.[194] The full parameters used in the molecular docking calculations using AutoDock Vina are provided in the supporting information.

Before the docking of our designed ligand to Mdm2 was explored, the docking of a known Mdm2 inhibitor, Nutlin-2, was explored.[195] This was performed to enable a comparison between the binding of our ligand to Mdm2 with that of a successful inhibitor. Although a crystal structure of Nutlin-2 and Mdm2 exists (PDB: 1RV1), the coordinates of Mdm2 used were still taken from its crystal structure with p53. A MM conformational search of Nutlin-2 was performed, and the lowest energy conformation found was used in the docking calculation. It is important to use the receptor from a different crystal structure to avoid biasing the results from re-docking.

Re-docking, when the ligand is docked into the receptor from the crystal structure it was extracted from, typically gives elevated binding affinities since the conformation of the binding site perfectly accommodates the ligand. Cross-docking however, when the ligand is docked into a receptor from a different crystal structure, avoids this problem.[253] Since a crystal structure of our designed ligand **96** with Mdm2 does not exist, the molecular docking calculation will always be a cross-docked simulation

and in order to compare our ligand with Nutlin-2, Nutlin-2 needs to be cross-docked with the same receptor as that chosen for our ligand i.e. from 1YCR.

The molecular docking calculation of Nutlin-2 with Mdm2 gave a predicted binding affinity of – 41.4 kJ mol^{-1}. The most favourable binding pose of Nutlin-2 found by AutoDock Vina closely reproduces the bound conformation of Nutlin-2 observed in the crystal structure (1RV1) (**Figure 57**). This confirms that the binding affinity of Nutlin-2 to Mdm2, predicted by AutoDock Vina, is reliable and can be used to compare the performance of our ligand too. The three binding groups of Nutlin-2 have been placed into the three hydrophobic binding pockets of Mdm2 and make extensive nonspecific van der Waals contacts with Mdm2. Slight conformational differences of these binding groups are however observed when compared to the bound conformation of Nutlin-2 in the crystal structure. In particular the 4-chlorophenyl group found in the Leu26 pocket. This could be caused by cross-docking Nutlin-2 with a different receptor. The crystal structure of Nutlin-2 with Mdm2 revealed that the 4-chlorophenyl group in the Leu26 pocket orients itself to engage in a π-π interaction with the His96 residue of Mdm2.[196] This π-π interaction will contribute to the overall conformation of the Mdm2 binding site which, as you can see clearly in **Figure 57**, is different to the conformation of the binding site that Nutlin-2 was docked into. By docking Nutlin-2 into Mdm2 with a different conformation of the binding site, this π-π interaction is lost and as a result the conformation of the 4-chlorophenyl group changes.

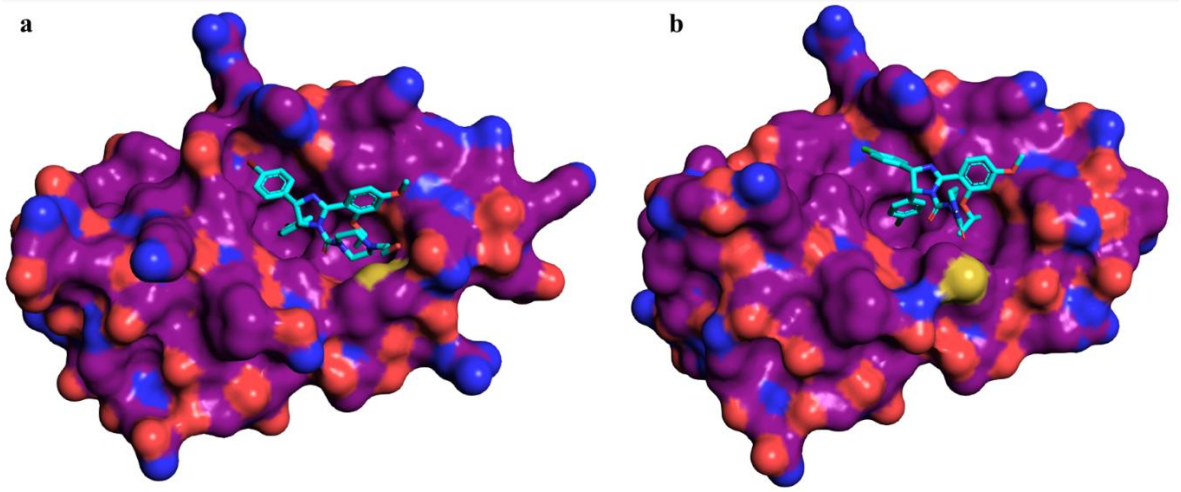


Figure 57. **a)** The crystal structure of Nutlin-2 and Mdm2 (PDB ID:1RV1) **b)** The binding pose of Nutlin-2 with Mdm2 found by AutoDock Vina.

The binding of our designed mimetic **96** to Mdm2 using AutoDock Vina was next explored. As mentioned, Mdm2 was taken from the crystal structure with p53 (1YCR) and the lowest energy conformation of **96** found from the MM conformational search was used in the calculation. The top binding pose of **96** found by AutoDock Vina places the hot-spot groups into the p53 binding pockets of Mdm2 (**Figure 58**). The three side chains have oriented themselves such that they mimic the three

binding groups of p53. However, the ligand is now binding in the opposite direction to what was expected by the design process i.e. the Leu mimetic is in the Phe19 pocket and vice versa. As discussed in the introduction, many small molecule inhibitors of the p53-Mdm2 PPI have been designed to place an aromatic group in the Leu26 pocket as it has been suggested that this improves the binding affinity due to π-π interactions with the His96 residue.[200] On examination of the binding of ligand **96** to Mdm2, predicted by AutoDock Vina, it is clear that a π-π interaction with His96 is unlikely to be occurring as it is not in the correct orientation to engage in this interaction, and so is not the reason as to why the **96** is binding in the opposite direction. It is also possible that the direction of the ligand is simply due to the fact that AutoDock Vina cannot differentiate between the size and the hydrophobic interactions made between the receptor and the Phenyl group vs the isopropyl group.

The predicted binding affinity between our designed mimetic **96** and Mdm2 is -36.8 kJ mol^{-1}. Although this is not as strong as Nutlin-2, it is a competitive binding affinity. The binding affinities obtained from AutoDock Vina typically range from -8.5 kJ mol^{-1} to -63 kJ mol^{-1}. Typically, a weakly binding inhibitor will have a binding affinity of <25 kJ mol^{-1}, a moderately binding inhibitor will have a binding affinity between 25-33 kJ mol^{-1} and a strongly binding inhibitor will have a binding affinity between 33-54 kJ mol^{-1}, while binding affinities >54 kJ mol^{-1} are only observed when reproducing a crystal structure (i.e. cross-docking).[242,254] Thus, with the binding affinity observed, **96** is expected to be a strongly binding inhibitor and capable of binding to Mdm2. The designed mimetic **96** is predicted to be a promising ligand for the inhibition of the p53-Mdm2 PPI and thus will be synthesised and analysed to confirm binding to Mdm2.

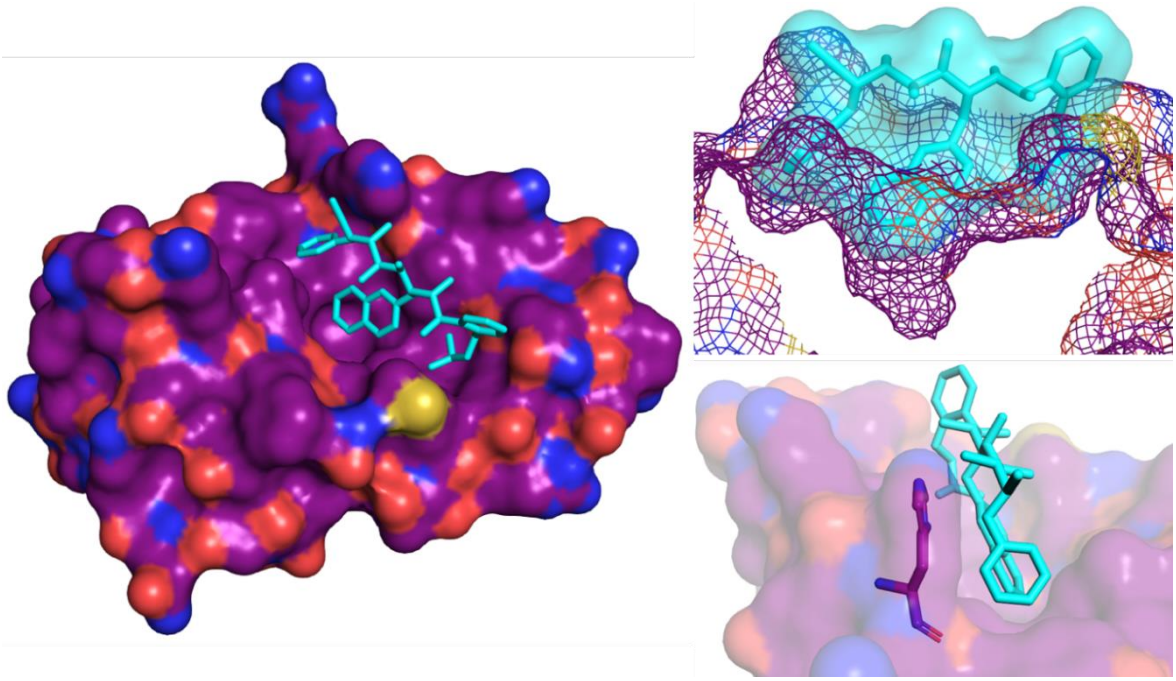


Figure 58. The predicted binding pose of **96** with Mdm2. It can clearly be seen that **96** is binding in the opposite direction as p53, with the benzyl group in the Leu26 pocket. The orientation of His96 of Mdm2 relative to **96** is also shown.

3.4 The Computational Design of a Library of p53 Mimetics

During SBDD, binding groups on the molecule will be altered in an attempt to increase the binding affinity between a ligand and its receptor through improved induced fit or by increasing the likelihood of interactions such as the hydrogen bond. The altered structures can then be tested computationally, using tools such as molecular docking, to investigate whether the redesigned molecule is a better candidate. This process was followed and a small library of compounds, differing in the hot-spot groups attached to the hydrocarbon backbone, was designed in an attempt to increase the binding affinity. The designed library of structures can be easily synthesised using lithiation–borylation by simply by substituting the appropriate benzoate ester into the sequence at judicious points.

Our first re-design was inspired by the work of Furet *et. al.* who have demonstrated that the correct placement of a chlorine atom on the group binding to the Trp23 pocket can significantly improve the binding affinity.[255] A chlorine atom is able to occupy a deep void in the Trp23 pocket and make additional van der Waals contacts with residues located deeper in the binding pocket, including Phe86 and Ile99.[256,257] It is for this reason that an overwhelming majority of both small molecule and peptide inhibitors of the p53-Mdm2 PPI contain a chlorine atom at a judicious position. Examining the binding of mimetic **96** to Mdm2, shown in **Figure 58**, it is clear that a large void below the naphthyl group in the Trp23 pocket is unoccupied. The effect of placing a chlorine atom at the 5- or 6-position of the naphthyl group (**103** and **104**) was explored. Using AutoDock Vina, the lowest energy conformer of both **103** and **104**, found from MM conformational search calculations, was docked into Mdm2. However, the hot-spot group was now too large to fit in the binding pocket and the void was not filled by the chlorine atom (**Figure 59**). As a result, the binding affinity was not improved.

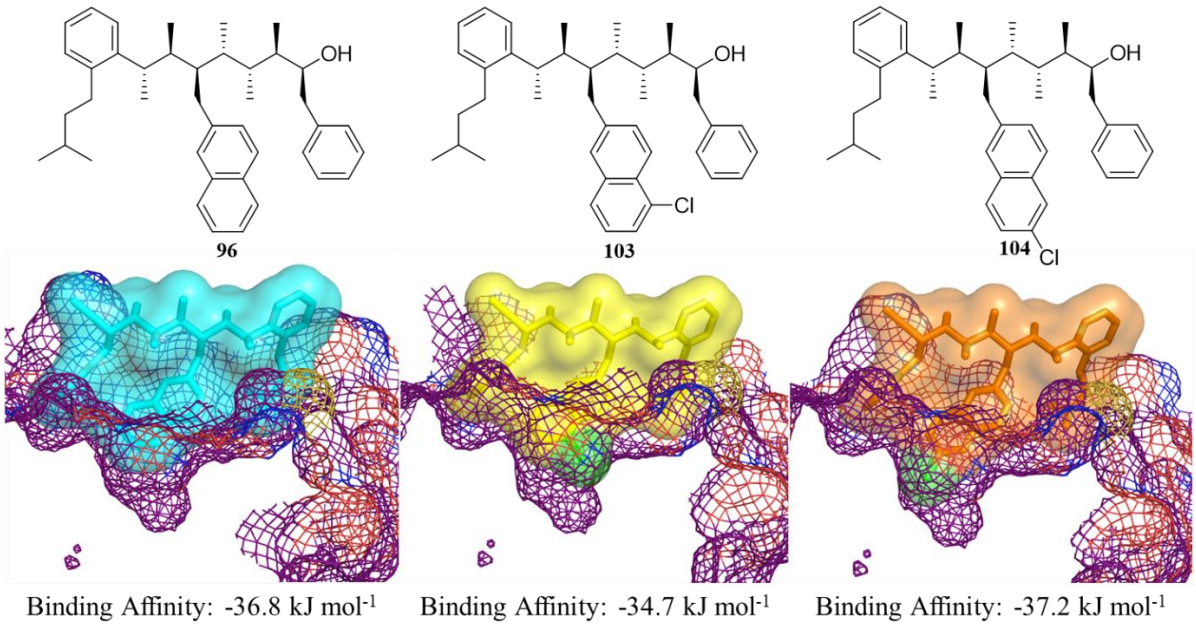


Figure 59. The binding of chloro-substituted naphthyl derivatives, assessed using AutoDock Vina.

This problem can be overcome by removing the chloro-substituted naphthyl group and replacing it with a *para*-chloro substituted phenyl group. To obtain the best induced fit, the chain length needs to be extended by one methylene unit. A conformational search of structure **105** revealed that a linear conformation was still preferred despite the additional chain length which increases the possibility of a *syn*-pentane interaction between the side chain and the backbone. The lowest energy conformation was subjected to molecular docking and a predicted binding affinity of -38.9 kJ mol^{-1} was observed (**Figure 60**). Examination of the binding pocket confirms that the chlorine atom has filled the void of the Trp23 pocket, and the increased binding affinity can be attributed to additional van der Waals contacts with Mdm2. It is worthwhile to note that the molecule has bound to Mdm2 in the opposite direction as **105**, with the phenyl group now in the Phe19 pocket and the isopropyl group now in the Leu26 pocket, as originally intended.

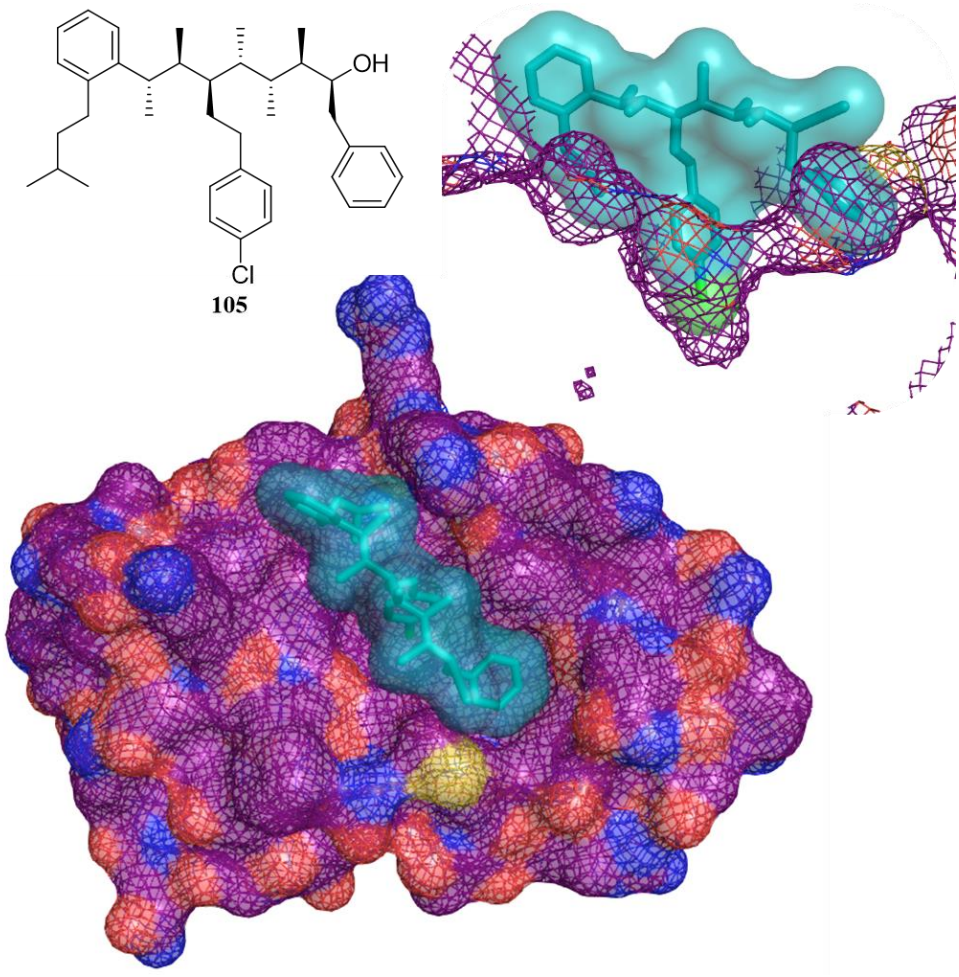


Figure 60. The predicted binding pose of candidate **105** as predicted by AutoDock Vina.

Examining the binding pocket when mimetic **105** is bound reveals that the Trp23 pocket still has space available. Vassilev has previously reported that the addition of fluorine atoms to the hydrophobic binding groups can improve the affinity of a class of pyrrolidine based Mdm2 inhibitors.[200] The addition of fluorine atoms to the two phenyl groups of **105** on the binding to Mdm2 was thus explored.

The fluorine atoms were added to the position on the ring deemed the most suitable for accommodation of the fluorine atom within the binding pocket. The resulting structure **106** bound to Mdm2 in the expected pose, with the binding groups located in the three binding pockets, in a similar manner to that observed with **105** (**Figure 61**). However, an increased binding affinity of -41 kJ mol^{-1} was obtained, attributed to increased van der Waals interactions from the two additional fluorine atoms with Mdm2.

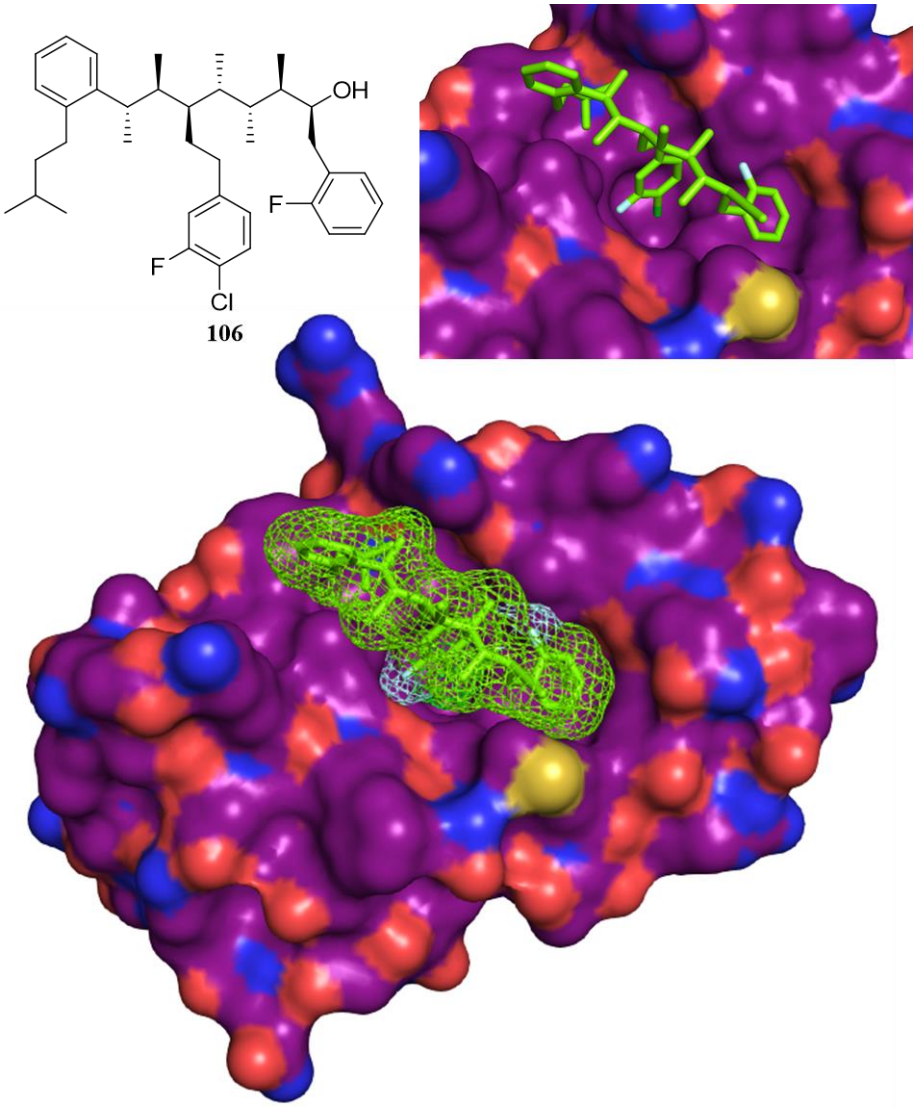


Figure 61. The predicted binding pose of candidate **106** as predicted by AutoDock Vina.

Although the majority of interactions between p53 and Mdm2 are hydrophobic, they are supplemented by a hydrogen bond between the indole NH of Trp23 on p53 and Leu54 on Mdm2. Many groups have reported that mimicking this hydrogen bond in a small molecule Mdm2 inhibitor is not important for strong inhibition of the p53-Mdm2 PPI however, AutoDock Vina was used to investigate the binding of an indole substituted mimetic (**107**) (**Figure 62**). Additionally, a chloro-substituted indole was investigated. As expected, the chloro-substituted indole (**108**) gave a stronger predicted binding affinity to Mdm2, compared to the indole derivative (**107**), as a result of increased van der Waals interactions.

However, in general the two indole substituted derivatives bind to Mdm2 with a lower affinity compared to the naphthyl- or phenyl- derivatives. On examination of the binding site with **107** and **108**, it is clear that the indole derivatives do not occupy the Mdm2 binding site to as great an extent as the other structures investigated. A further caveat of the indole derivatives is that their synthesis will be more challenging owing to the NH of the indole which will require protection if it is to survive the lithiation–borylation sequence.

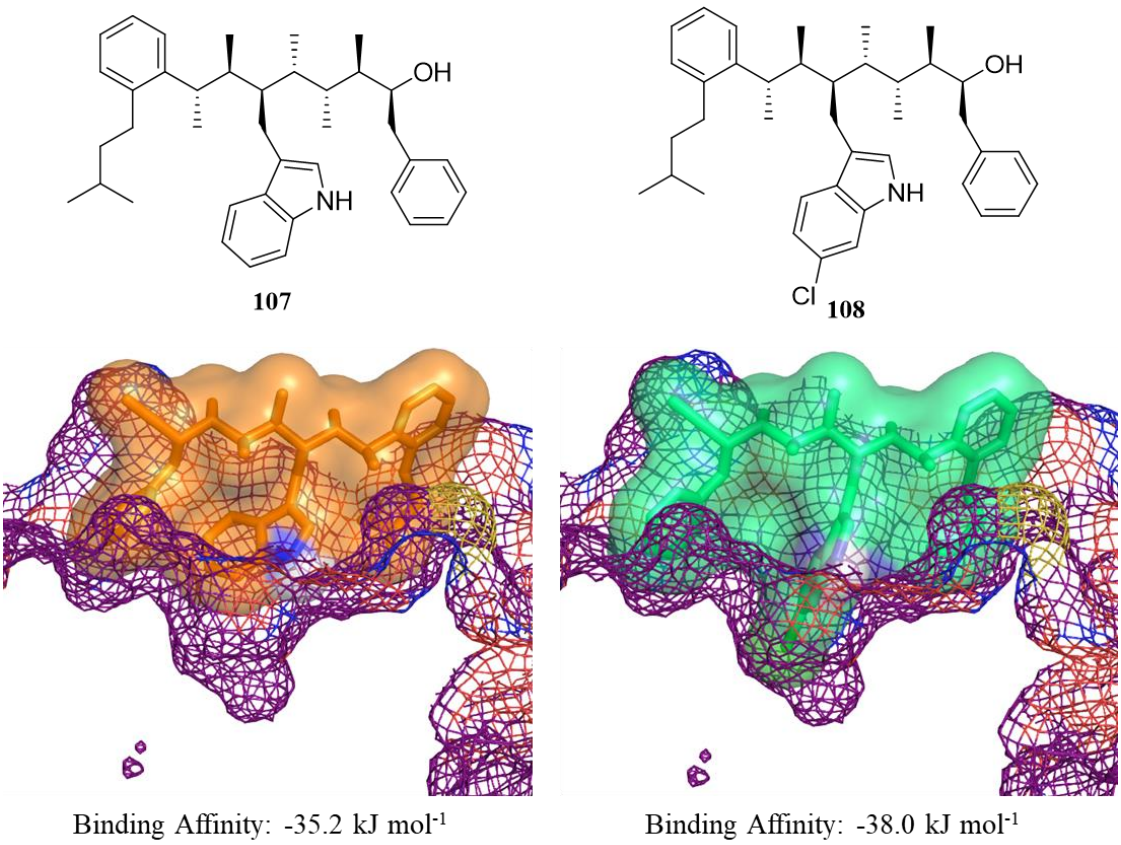


Figure 62. The predicted binding poses of indole derivatives **107** and **108** as predicted by AutoDock Vina.

A final structure (**110**) investigated is shown in **Figure 63**. An oxygen atom has replaced a CH_2 group in the side chain for the Trp23 hot-spot group mimetic. Recently Bootwicha *et. al.* have shown that the assembly line synthesis of polypropionate motif through the iterative introduction of carbinol units is possible. Using enantioenriched lithiated α-chlorosilanes in the reagent-controlled homologation of boronic esters yields α-silylalkyl boronates with exceptional stereocontrol. Oxidation of the silyl group reveals the corresponding alcohol.[258] Using this methodology interspersed with the previously described methods for homologating boronic esters using benzoate esters, mimetic **109** could be synthesised. From this derivative, alkylation of the alcohol could reveal a number of possible mimetics. The binding of a possible mimetic (**110**) was explored by AutoDock Vina. The MM conformational search calculations revealed that the presence of the oxygen atom did not significantly affect the conformational bias with a linear conformation still preferred, with 67% of conformers being linear.

The lowest energy conformer of **110** was docked into Mdm2, which revealed that the incorporation of the oxygen atom in the side chain does not significantly diminish the binding to Mdm2 and a binding affinity of - 38.1 kJ mol^{-1} was obtained.

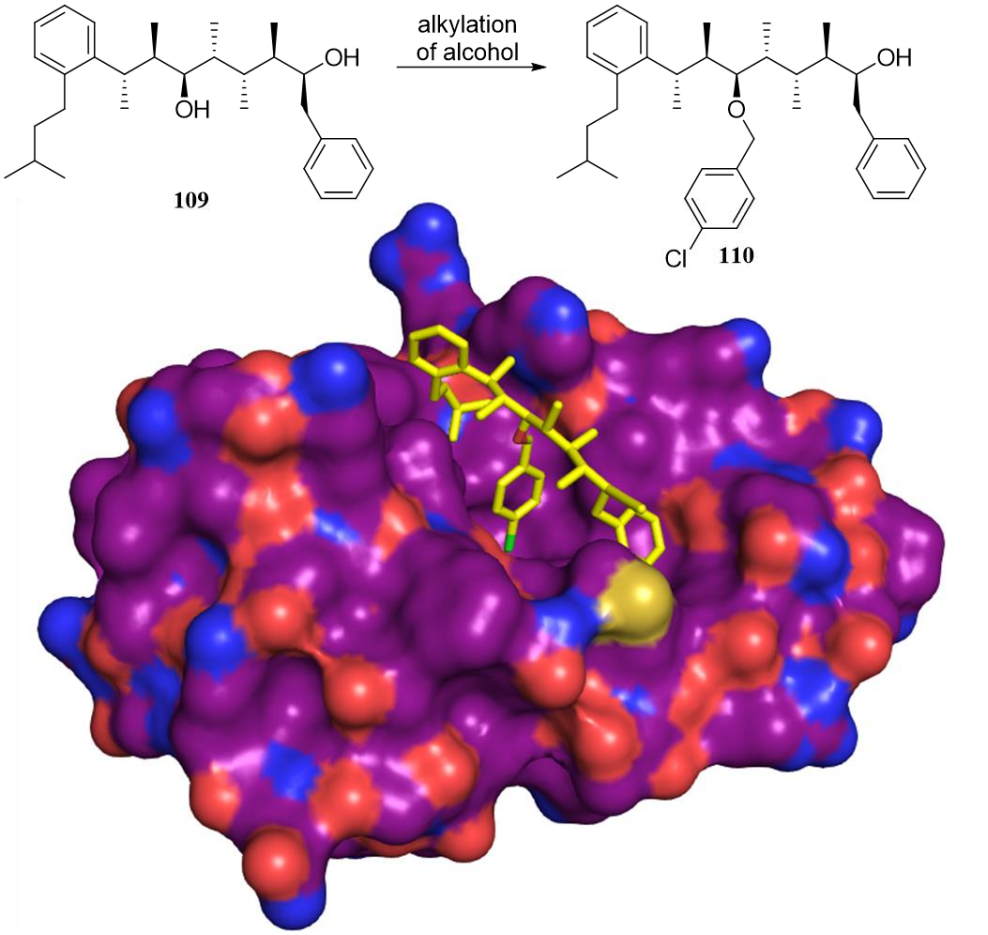


Figure 63. The predicted binding pose of candidate **110** as predicted by AutoDock Vina.

In addition to designing a library of structures that will bind well to Mdm2, the importance of both the hot-spot groups and the conformational control in obtaining high affinity candidates in the inhibition of PPIs needs to be confirmed. The importance of the hot-spot groups in PPI inhibition is well known and many groups have reported weaker binding affinities if the hot-spot groups are removed from the inhibitors.[203] This will be confirmed experimentally, by synthesising the all methyl substituted structure **111**. This candidate still exhibits excellent conformational control with over 97% of conformers adopting a linear conformation, calculated by a MM conformational search, however a weaker binding affinity due to the removal of the hot-spot groups is expected. Molecular docking of the lowest energy conformer of **111** revealed a weaker predicted binding affinity of -29.0 kJ mol^{-1}. It is clear from **Figure 64** that the Trp23 binding pocket is not occupied to any extent with mimetic **111**, owing to the weaker binding affinity.

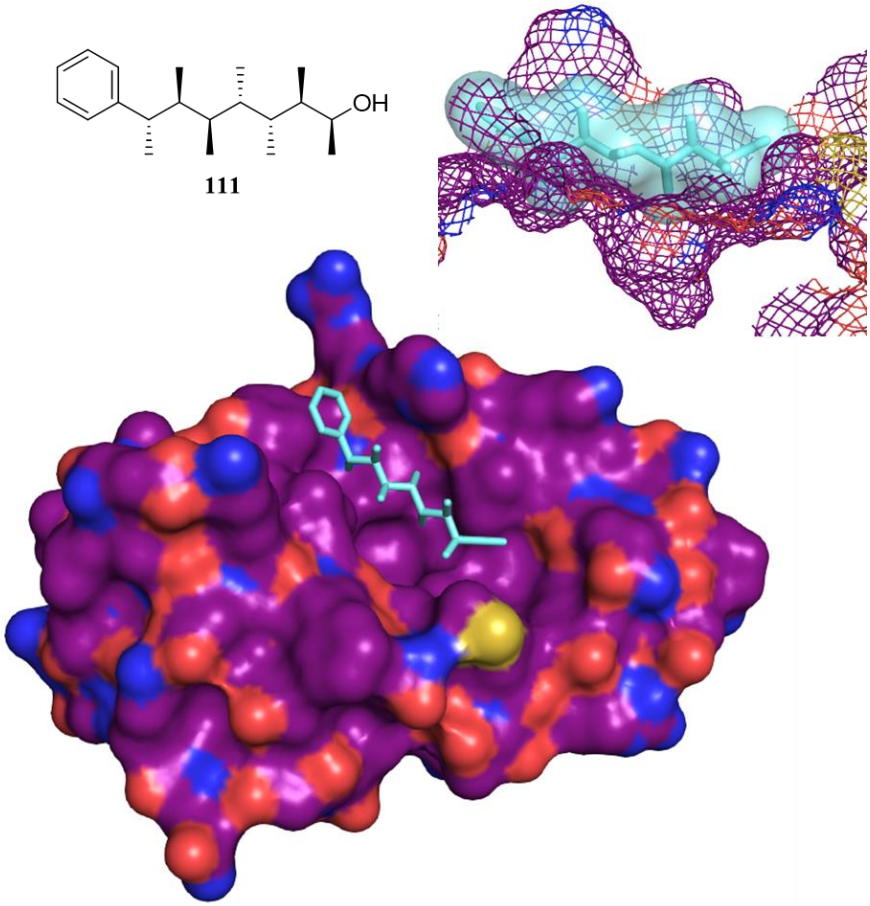


Figure 64. The predicted binding pose of control candidate **111** as predicted by AutoDock Vina.

The importance of conformational control in PPI inhibition can be confirmed by removing the methyl groups from structure **96** and leaving only the hot-spot groups, resulting in structure **112**. The conformation of the molecule will no longer be controlled by the avoidance of *syn*-pentane interactions and no conformational bias of the molecule is expected. Indeed, a MM conformational search of structure **112** revealed no conformational bias of the molecule with only 1.4% of conformers under 21 kJ mol^{-1} adopting a linear conformation. The first conformer to adopt a linear conformation is 6.23 kJ mol^{-1} above the global energy minima found by MM. Thus, a significant conformational reorganisation is required to access the linear conformer, which represents the expected bound conformation.

Previously, molecular docking was performed on the lowest energy conformer of the candidates from a MM conformational search, which represented the overwhelming conformational bias observed. For candidate **112**, no conformational bias was observed and thus no clear conformation could be taken through to molecular docking (**Figure 65**). It is expected that a weaker binding affinity will be observed due to the conformational reorganisation that will need to take place to access the bioactive conformation. This energy penalty cannot be evaluated using molecular docking and instead much more sophisticated computational methods, such as molecular dynamics, would be needed to provide an accurate estimation of the binding affinity for candidate **112**. Thus, the binding affinity for candidate

112 could not be explored using molecular docking, however it will be tested experimentally to confirm the importance of conformational control.

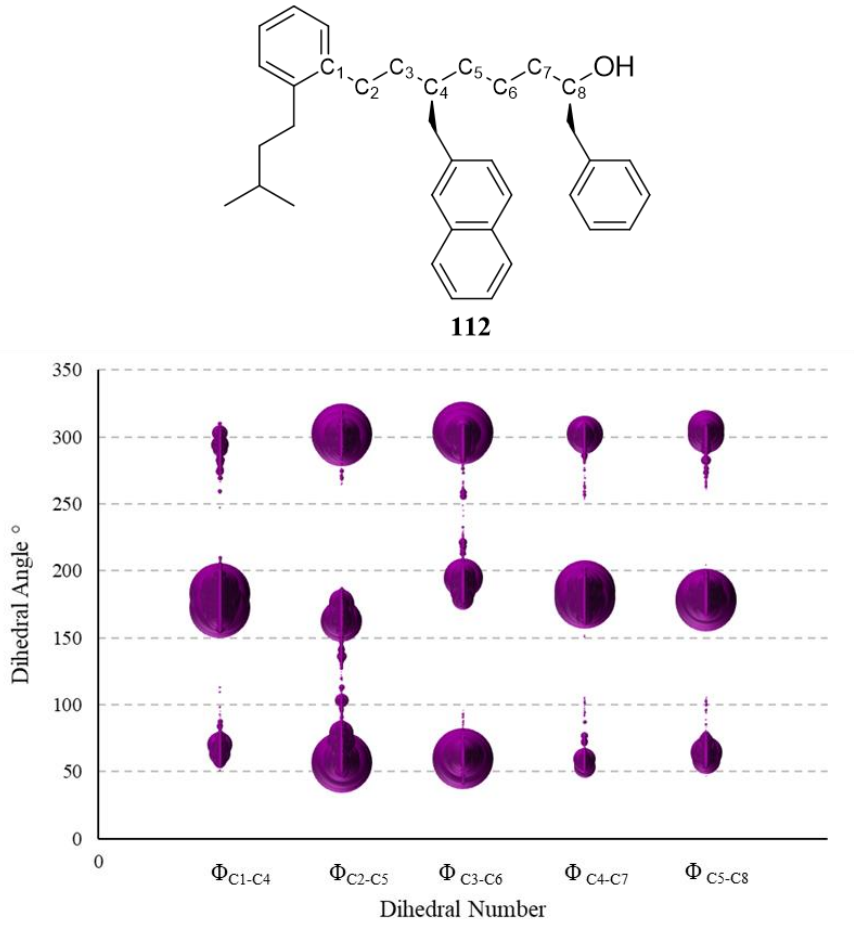


Figure 65. The MM conformational search results of control molecule **112** which reveals no conformational bias is observed for the molecule.

Using both MM conformational searching and molecular docking using AutoDock Vina a small library of potential Mdm2 inhibitors has been designed. MM conformational searching has confirmed that the addition of hot-spot groups to the scaffold does not perturb the conformational bias, with a linear conformation preferred for all structures (with the exemption of structure **112**, which was designed to display no conformational control). The estimated binding affinities for the candidates reveal strongly binding structures with candidate **106** displaying a binding affinity competitive with that of one of the most successful inhibitors of the p53-Mdm2 PPI, Nutlin-2. The synthesis of some of these structures using lithiation–borylation will be discussed in the next chapter.

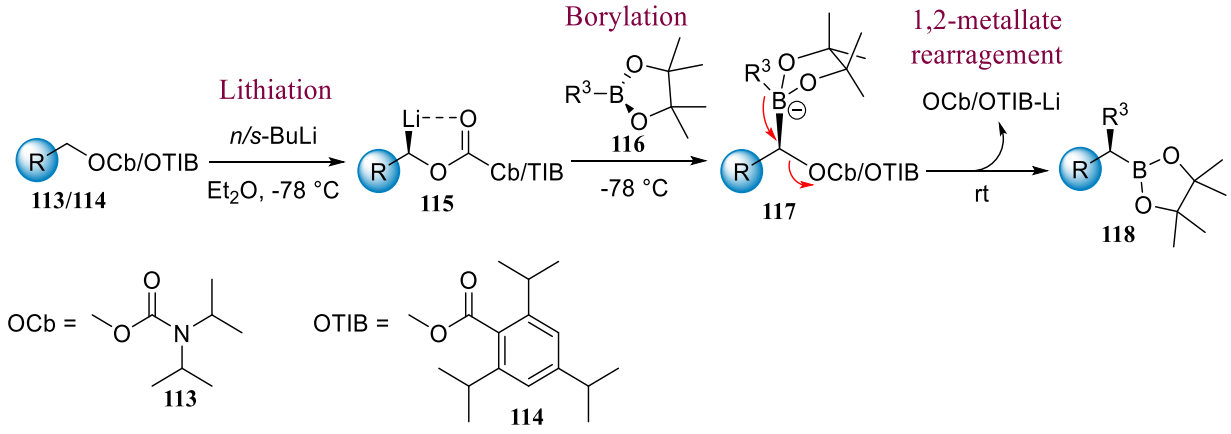
Chapter 4: The Synthesis of an α-Helix Mimetic of p53

The methods developed and molecules synthesised in Section 4.1 and 4.2 were performed by Lydia Dewis. The molecules synthesised in Section 4.3 and 4.4 were performed by Dr. Madhavachary Rudrakshula.

In order to experimentally confirm that the designed p53 mimetics adopt the conformation expected and can bind to Mdm2, they need to be synthesised. As discussed in section 1.2 of chapter 1, the hydrocarbon scaffold is accessible using lithiation–borylation. This methodology allows hot-spot mimetic groups to be substituted onto the scaffold in high yields and with high diasteroselectivities. Using lithiation–borylation, a variety of hot-spot mimetic groups can be incorporated onto the scaffold, simply by the substitution of the appropriate benzoate ester into the sequence at judicious points. Thus, using only one synthetic methodology one can access a small molecule α-helical scaffold bearing the appropriate hot-spot groups to mimic the α-helix in question.

4.1 Using Lithiation–Borylation to Synthesise α-Helix Mimetics

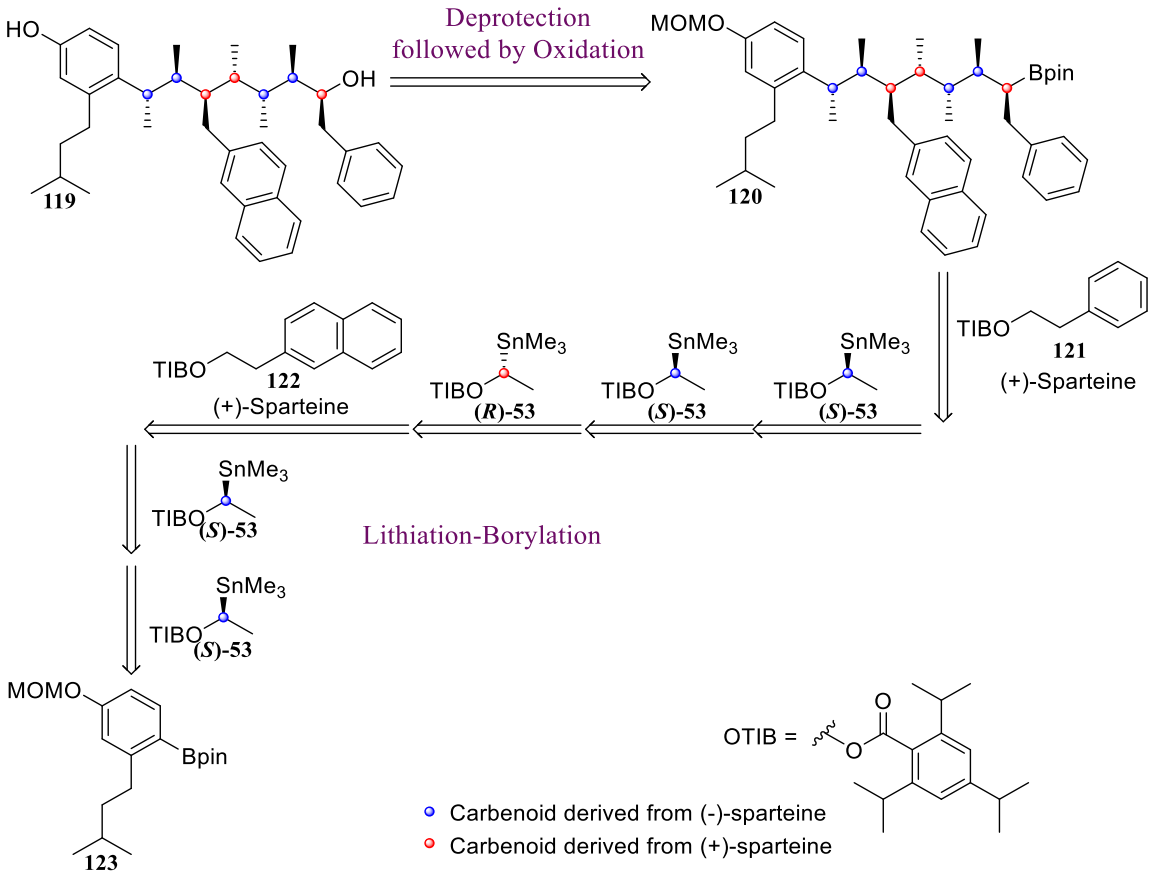
The homologation of boronic esters is a powerful C–C bond forming technique that typically displays high yields and if required high stereoselectivities. This process proceeds through three main steps: 1) Enantioselective lithiation of a carbenoid precursor (**113/114**) to generate a chiral carbenoid (**115**) that is chemically and configurationally stable at –78 °C. 2) Borylation, during which the chiral carbenoid is trapped by an electrophilic boronic ester species (**116**) to yield a boronate complex (**117**). This is also chemically and configurationally stable at –78 °C. 3) A stereospecific 1,2-metallate rearrangement upon warming the boronate complex to room temperature. This forms a new C–C bond to yield a homologated boronic ester (**118**) that is ready for subsequent homologations (**Scheme 9**).



Scheme 9. The general mechanism of lithiation–borylation.

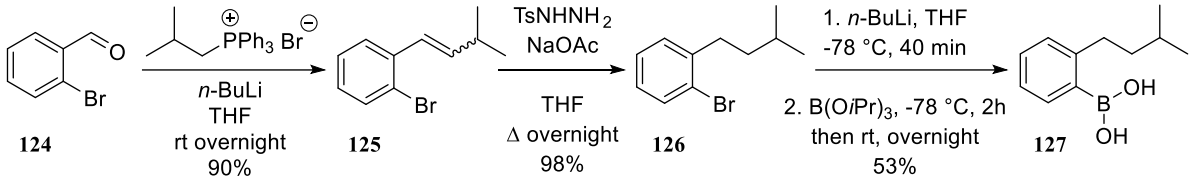
The proposed retrosynthesis of a phenol derivative **119** of the first designed p53 mimetic, **96**, is shown in **Scheme 10**. The phenol group has been incorporated to provide a handle to allow for functionalisation

to improve the scaffold aqueous solubility. The phenol group will be revealed from a deprotection of the MOM ether. This will be followed by the oxidation of the terminal boronic ester to the secondary alcohol which can also be further functionalised to improve solubility. More will be discussed on strategies to improve aqueous solubility in section 4.2. The retrosynthetic analysis of molecule **119** reveals a total of seven homologations to be performed using the lithiation–borylation methodology. Different strategies to generate the chiral carbenoid will be used depending on whether a methyl group or a hot-spot group is attached. Previously, Aggarwal *et al*. have shown that tin-lithium exchange of enantioenriched stannanes can be used to generate the chiral lithium carbenoid, and this carbenoid can be used in an iterative lithiation–borylation sequence since the required stannane is crystalline and can be recrystallised to >99.9:0.1 *e.r.*[27] For the synthesis of **119**, the carbenoid precursor used in the sequence will change from enantioenriched stannanes to primary benzoates, depending on the group being installed onto the growing carbon chain. Using the chiral ligand-assisted deprotonation of primary benzoates to generate the chiral carbenoid results in a slight decrease in the diastereoselectivity of the reaction and as a result, column chromatography will be performed at certain points in the synthesis. Additionally, the boronic ester being homologated in this synthesis will become sterically very hindered, which can decrease the rate of borylation, and potentially lead to under-homologated product, which is important to remove by purification.



Scheme 10. The proposed retrosynthesis of the designed p53 mimetic, **119**.

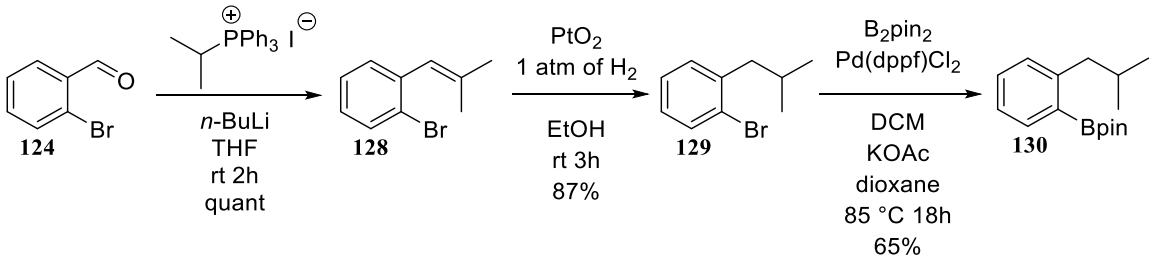
Prior to starting the lithiation–borylation sequence, the starting boronic ester (**123**) needs to be synthesised. The synthesis of **123** was based on two previous syntheses of similar organoboron species. In 2006, Spencer *et. al.* reported the synthesis of photoactivatable analogues of cholesterol involving a Suzuki carbonylative coupling with the phenylboronic acid **127** (**Scheme 11**).[259] The synthesis reported by Spencer is high yielding and involves reactions that are well reported and known to be widely applicable.



Scheme 11. The synthesis of boronic acid **127** by Spencer *et. al.* in 2006.[259]

It is interesting to note the choice of conditions for the hydrogenation. Typically, hydrogenation is performed using heterogeneous or homogeneous catalysis using gaseous hydrogen and a metal catalyst, ordinarily palladium. The diimide reduction employed by Spencer is not frequently used however, it does provide a metal-free alternative to catalytic hydrogenation reactions that avoids the use of gaseous hydrogen. The diimide reduction is also advantageous if the substrate in question contains sensitive O-O or N-O bonds since it does not lead to reductive cleavage of these bonds, a common problem when using catalytic hydrogenation conditions. Diimide reductions do however typically suffer from long reactions times and high reaction temperatures.[260] Additionally, a large excess of the reagents that generate the diimide are required to compete with both the decomposition of diimide to N_2 and H_2 and the disproportionation of diimide to hydrazine and N_2.[261]

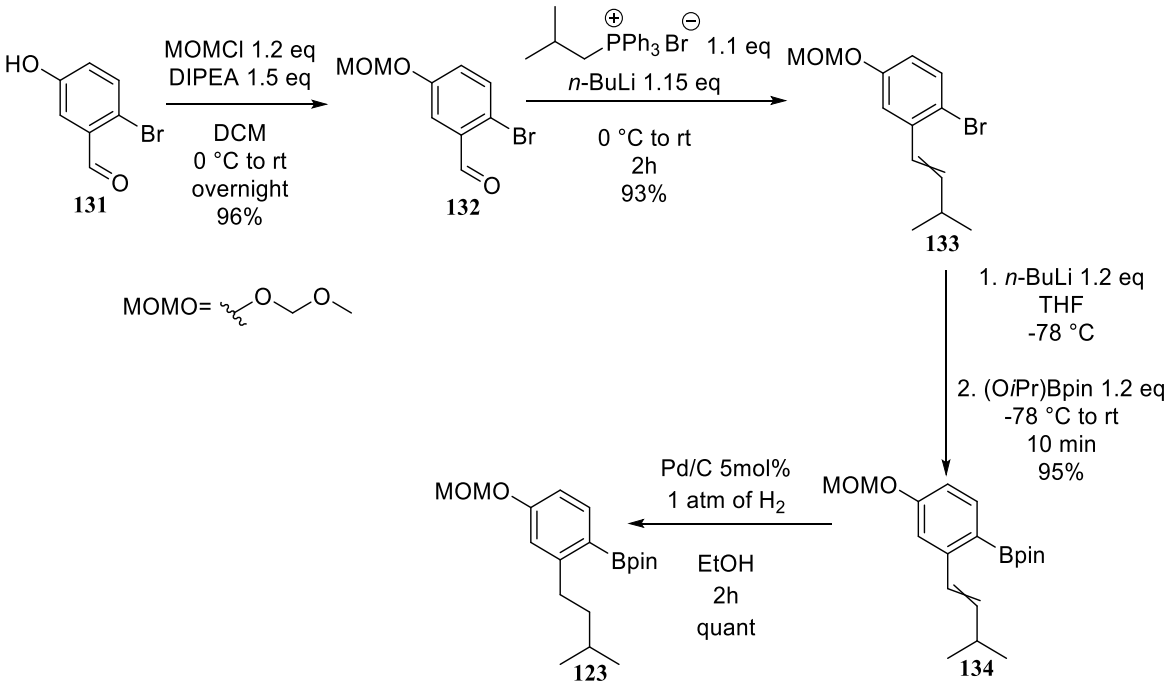
A second reported synthesis of an organoboron species similar to the starting boronic ester **123** was reported in 2016 by Tuck *et. al.* (**Scheme 12**).[262] The boronic ester **130** was a required coupling partner in a Suzuki cross-coupling reaction to synthesise a pyrimidine-based tercyclic α-helix mimetic. The first step involves a Wittig reaction, similar to the one reported by Spencer, followed by a hydrogenation using more typical platinum catalysis conditions. The resulting aryl halide **129** was subjected to a Miyaura borylation to afford the final aryl boronic ester **130**.



Scheme 12. The synthesis of boronic ester **130** by Tuck *et. al.* in 2016.[262]

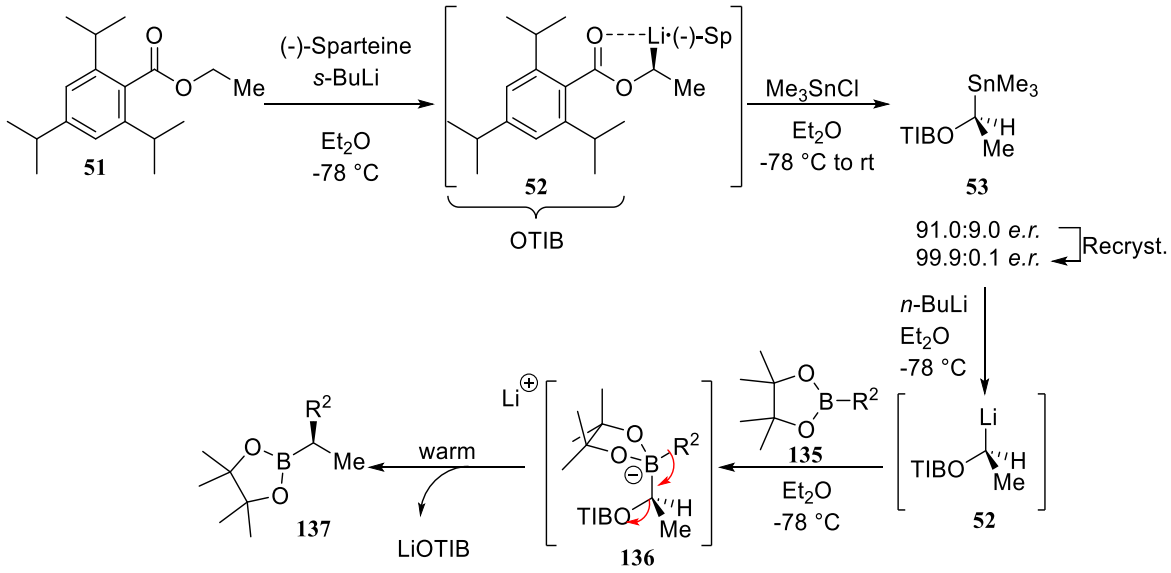
Taking inspiration from the two previous syntheses described, we report the following synthesis for the starting boronic ester **123** (**Scheme 13**). The first step involves a MOM protection of the commercially available phenol **131**. The MOM protection of phenol **131** has been described previously in the literature, with a yield of 76% reported.[263] Using the same conditions as described previously, a higher yield of 96% was obtained in the synthesis of **132**. This yield is reproducible on both a small (<100 mg) and moderate (>5g) scale. Following the MOM protection of phenol **131**, the Wittig reaction was performed using *iso*butyltriphenylphosphonium bromide providing a high yield of 93% of **133**, comparable to the yields observed in the Wittig reactions performed by Tuck and Spencer. Again, this yield is reproducible on both a small and moderate scale.

After the installation of the first hot-spot group by the Wittig reaction, halogen-lithium exchange followed by borylation using 2-isopropoxy pinacolborane ((O^{i}Pr)Bpin) was performed on **133** to install the pinacol boronic ester needed to start the lithiation–borylation sequence. This two-step reaction yielded the boronic ester **134** in a 95% yield, and the two-step sequence could be performed on both a small and moderate scale with no change in yield. The final step towards the synthesis of the starting boronic ester **123**, was a palladium catalysed hydrogenation of the alkene to the corresponding alkane. Using 5% palladium on charcoal (Pd/C) with a H_2 balloon in ethanol proved successful, and the hydrogenation was completed in quantitative yield. When this reaction was performed on a small scale, catalyst loadings of 10% were employed. However, upon increasing the scale of the reaction to >5g, catalyst loadings could be reduced to 5%.



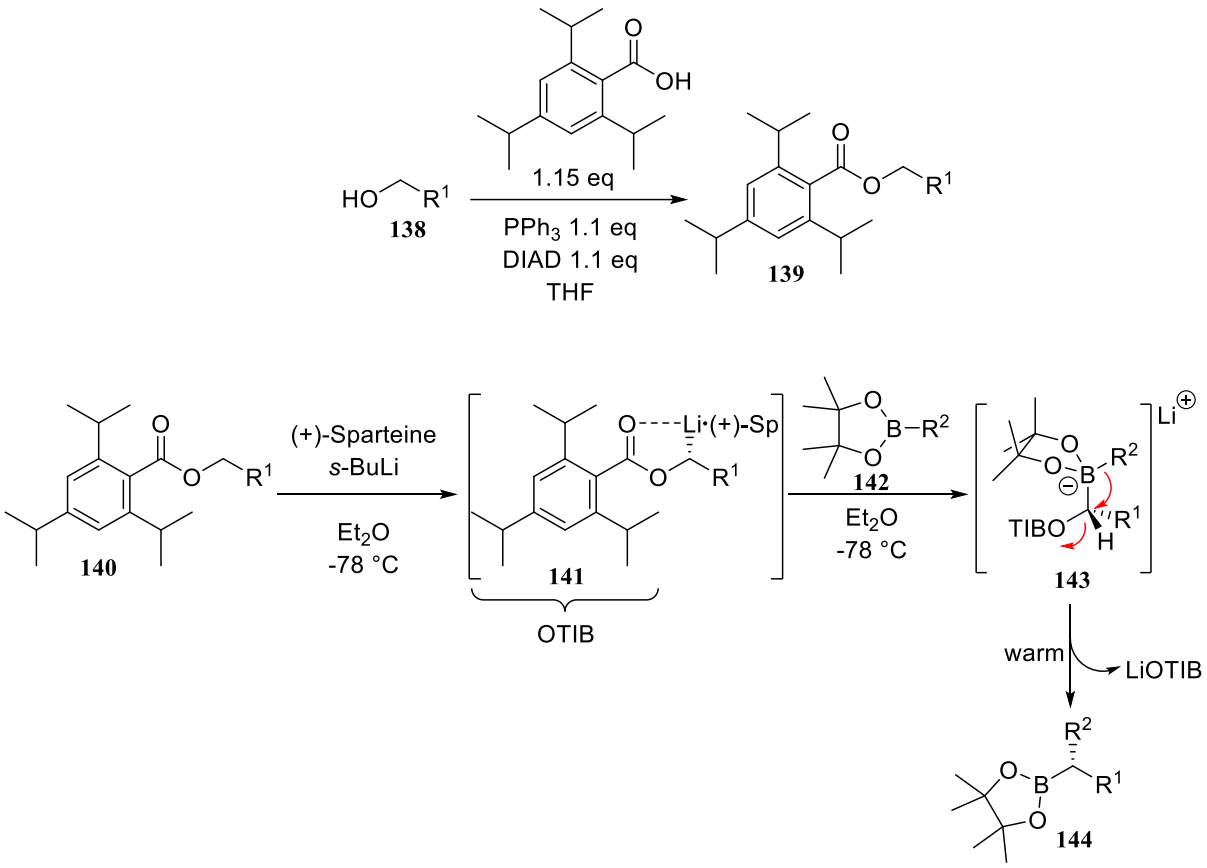
Scheme 13. The synthesis of starting boronic ester **123.**

The next stage of the synthesis of the designed p53 mimetic **119** involves the homologation of boronic ester **123**. To install the methyl groups, the methodology developed by Aggarwal *et. al.* was employed during which enantioenriched stannanes are used to generate the chiral carbenoid.[260] The two enantiomers of the stannane required in the synthesis of **119** can be generated from the deprotonation of ethyl 2,4,6-triisopropylbenzoate **51** using *s*-BuLi in the presence of (+)-sparteine or (-)-sparteine, to generate the carbenoid **52**. This was followed by electrophilic trapping with trimethyltinchloride to generate stannane **53** in a 91.0:9.0 enantiomeric ratio. Stannane **53** is a white crystalline solid and can be recrystallised in methanol to >99.9:0.1 *e.r* (**Scheme 14**). The stannanes are bench stable and can be stored until needed. Upon tin-lithium exchange of the enantioenriched stannanes with *n*-BuLi at –78 °C, the corresponding chiral carbenoid is formed which can be used to homologate the desired boronic ester. Using this methodology, both enantiomers of the stannane **53** can be prepared on a large scale (>10g), recrystallised to >99.9:0.1 *e.r* and simply stored under air.



Scheme 14. The generation of enantioenriched stannanes and the corresponding stereospecific tin-lithium exchange to generate the chiral carbenoid needed for the homologation of boronic esters.

This methodology was not performed when installing the hot-spot groups for several reasons. The first being that the corresponding stannane of the benzoate ester needed to install appropriate hot-spot groups will most likely not be a crystalline solid. The enantioenriched stannanes need to be recrystallised if they are to be accessed in the high enantiopurities required for lithiation–borylation. Secondly, the synthesis of the enantioenriched stannanes is much more effective on a large scale, principally to increase the percentage recovery of the sparteine used in their synthesis. Since the hot-spot groups are only introduced once in the sequence, synthesising large amounts is unnecessary. The hot-spot groups will therefore be installed using the sparteine assisted deprotonation of the corresponding benzoate ester. The benzoate esters, **139**, can be synthesised by Mitsunobu reaction of the corresponding alcohols, **138**, which were both commercially available (**Scheme 15**).



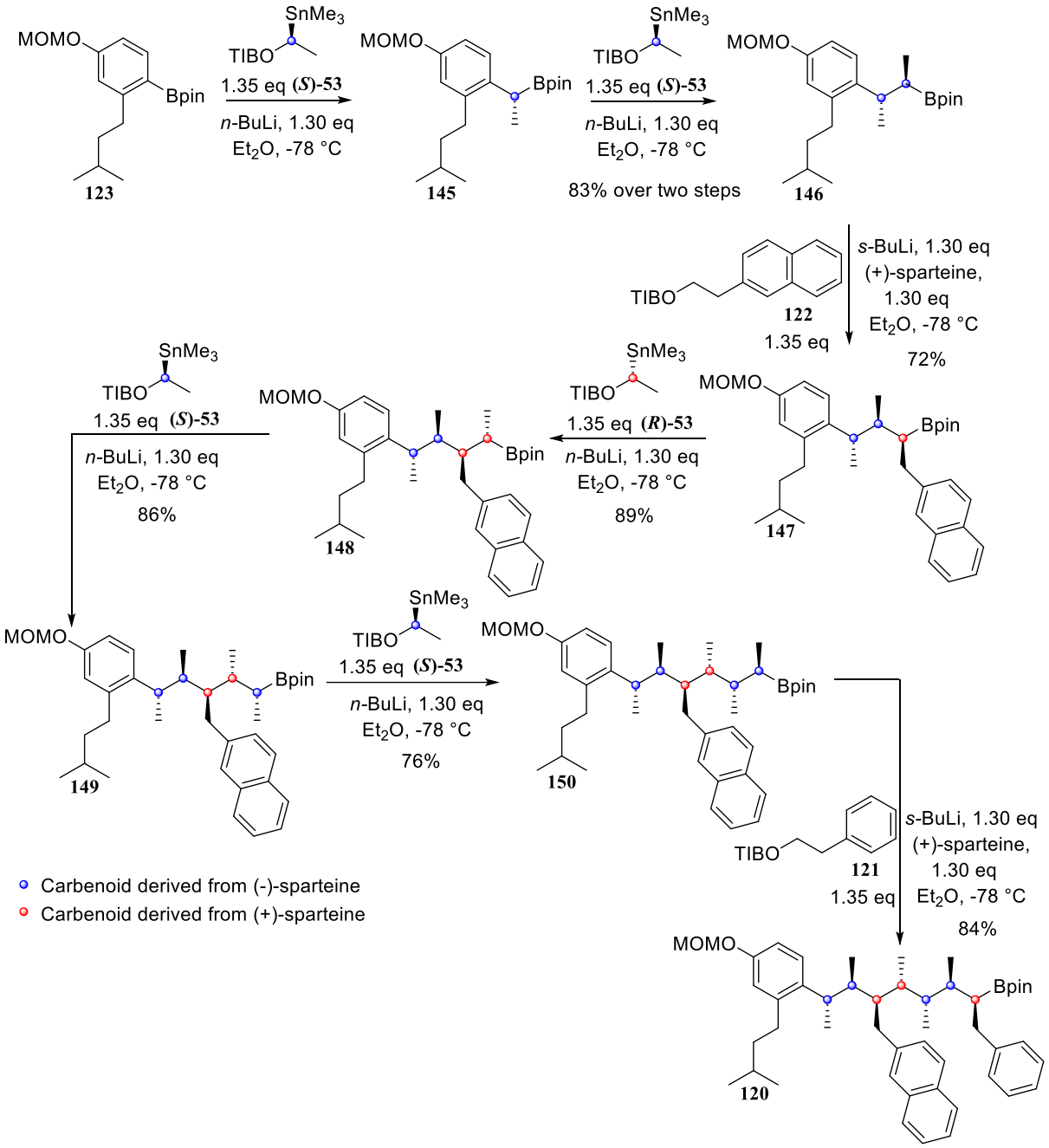
Scheme 15. The Misunobu reaction to form the TIB esters from the corresponding alcohols followed by sparteine mediated asymmetric deprotonation of a benzoate ester in the homologation of a boronic ester.

An optimised protocol for the iterative homologation of boronic esters using enantioenriched stannanes has been established by Aggarwal *et. al.* Nine consecutive homologations were performed with an aqueous workup performed after every third homologation, with column chromatography performed after the final homologation only. The fact that column chromatography was not required until the final homologation was only possible due to the very high enantiomeric ratio that the stannanes can be accessed in and the impressive conversions of >95% for each homologation. Since the sparteine mediated asymmetric deprotonation of benzoates is also being used in the synthesis of **119**, which typically occurs with lower enantiomeric ratios of ~95.0:5.0, each homologation cannot be performed iteratively.

In total seven homologations, starting from boronic ester **123** were performed, to yield the final boronic ester **120** (**Scheme 16**). The first two homologations were performed iteratively following the exact procedure reported by Aggarwal *et. al.*[27] Following the first two iterative homologations using enantioenriched stannanes, the second hot-spot group was installed by sparteine-mediated homologation. The lithiation time for benzoate **122** can be estimated to be ~1 hr, based on an extensive study into the reaction times of lithiation and borylation using different benzoates, carbamates, diamines and reaction conditions, by *in situ* IR spectroscopy by Aggarwal *et. al.*[78] Similarly, this study can be

used to estimate a borylation time with boronic ester **122** of <3 hrs. The 1,2-metallate rearrangement can be monitored using ^{11}B NMR, which revealed that full migration had occurred after 2 hrs. Complete conversion of boronic ester **146** to boronic ester **147** was revealed by TLC analysis. The crude reaction mixture was purified by column chromatography to remove any unreacted TIB ester **122**, which is present in excess, and to remove any under-homologated product. Following column chromatography, boronic ester **147** was obtained in 72% yield. The ^{1}H and ^{13}C NMR spectra of **147** revealed one diastereomer with a *d.r.* of >95:5.

A further three homologations were performed using enantioenriched stannanes, with borylation times of 3 hrs to ensure complete conversion of the lithiated carbenoid to the boronate complex. The 1,2-migration was followed by ^{11}B NMR, which revealed full migration was achieved for all three homologations after 2 hrs. Column chromatography was performed after each homologation to remove any under-homologated product. ^{1}H and ^{13}C NMR revealed one diastereomer with a *d.r.* of >95:5. A final sparteine mediated homologation of boronic ester **150** was performed using the benzoate ester **121** to install the final hot-spot group. The lithiation of this benzoate ester has been reported previously by Aggarwal *et. al.* and full lithiation was achieved within 1 hr.[78] Following lithiation of the benzoate ester **121**, boronic ester **150** was added to the reaction mixture and the reaction mixture was stirred at - 78 °C for 3 hrs. The 1,2-metallate rearrangement was again followed by ^{11}B NMR, and after 2 hrs full migration had occurred. The final boronic ester **120** was accessed in 84% yield, following column chromatography. Each homologation proceeded in high yields, to give a yield of 29% over the seven homologations.



Scheme 16. The forward synthesis of boronic ester **120** using the homologation of boronic esters.

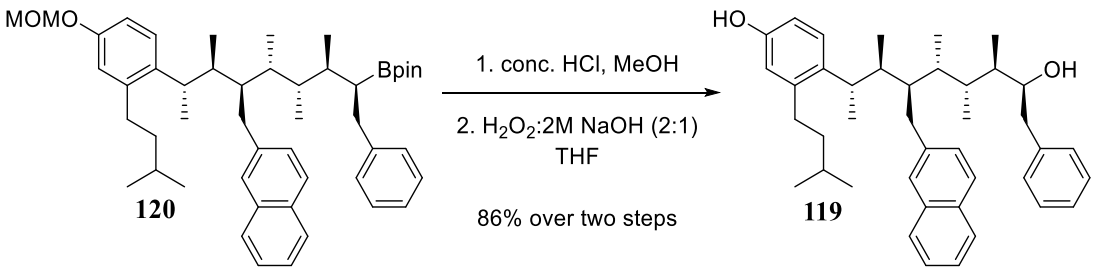
The aim of this project is to demonstrate that the conformationally controlled α-helix mimetics designed and synthesised can bind to the p53 binding domain of Mdm2. In order to confirm that the designed p53 mimetic binds to the correct location on Mdm2, experimental procedures to measure protein-ligand binding will need to be performed. However, since Mdm2 is only stable in aqueous environments, the designed p53 mimetics will need to be functionalised to improve their water solubility as they are currently very lipophilic. Fortunately, the boronic ester is one of the most useful groups for further functionalisation and can be easily transformed into many other functional groups. The next section will discuss the functionalisation of boronic ester **120** for improved water solubility.

4.2 Improving the Aqueous Solubility of the p53 Mimetic

The current designed mimetic **120** is very lipophilic and will display poor water solubility. Poor water solubility is a very common issue with small molecule inhibitors of PPIs, since the binding interactions between the proteins engaged in a PPI are usually very hydrophobic, thus any small molecule capable of interrupting this interaction needs to contain a large number of hydrophobic groups. Therefore, in order to experimentally confirm binding of our designed mimetic to Mdm2, water solubilising groups need to be installed onto the scaffold. It is hoped that the addition of hydrophilic groups at either ends of the molecule will improve its aqueous solubility whilst maintaining its conformational preference and ability to bind to Mdm2.

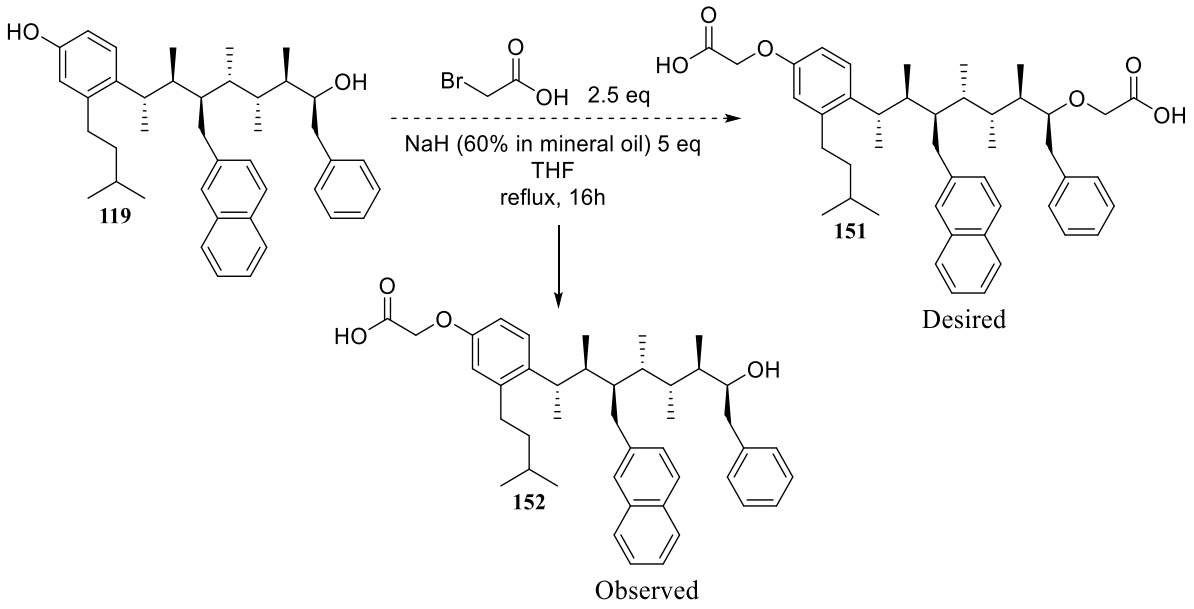
When it comes to choosing how to functionalise the terminal boronic ester there are a large number of choices. The terminal boronic ester can be oxidised to reveal the alcohol which can be further functionalised. Alternatively, amination can be performed to transform the boronic ester to the corresponding amine which again can be further functionalised. Aggarwal and co-workers have also shown that boronic esters can also be converted to alkyne derivatives[264] which can be reacted with azides bearing water solubilising groups to form the corresponding triazole in the 'click' reaction.

The terphenyl scaffold synthesised by Hamilton *et. al.* were used as a starting point for deciding the best direction for functionalisation.[265] The terphenyl molecules are also very lipophilic and Hamilton *et. al.* improved water solubility by functionalising with terminal carboxylic acids located at either end of the terphenyl scaffold. Taking inspiration from Hamilton *et. al.* the addition of terminal carboxylic acid groups to the scaffold was explored. Terminal carboxylic acid groups can be installed by first revealing the phenol, from deprotection of the MOM ether, followed by oxidation of the boronic ester to yield the secondary alcohol (**Scheme 17**). The phenol and alcohol could then be simultaneously deprotonated with a suitable base and subsequently alkylated with appropriate groups to install terminal carboxylic acids on either end of the scaffold. The deprotection of the phenol and oxidation of the boronic ester were both high yielding with a yield of 86% obtained over the two steps, to reveal the bisalcohol **119**.



Scheme 17. The deprotection of the MOM ether followed by oxidation of the boronic ester to yield the bisalcohol **119**

The next step involves the deprotonation of both the phenol and the secondary alcohol, followed by alkylation. The initial conditions employed in the alkylation are shown in **Scheme 18**. These conditions have been used routinely in the literature for the alkylation of both phenols and secondary or tertiary alcohols.[266,267] Unfortunately, only alkylation of the phenol was achieved (**152**), confirmed by NMR. The secondary alcohol is likely to be very sterically hindered making it a poor nucleophile, explaining the monoalkylation observed.



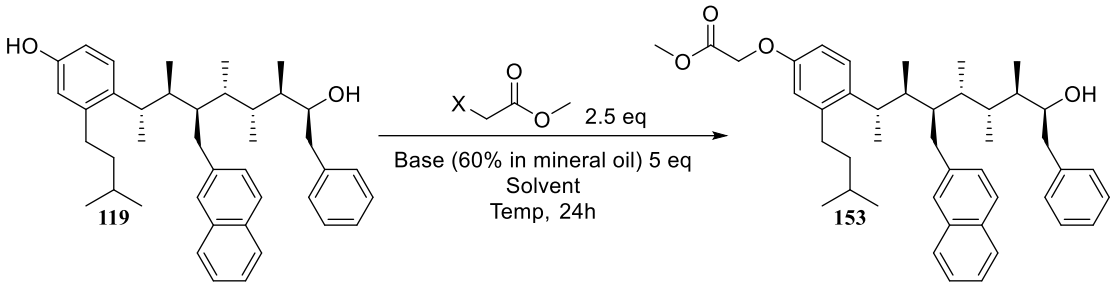
Scheme 18. The attempted alkylation of the phenol and the alcohol to install two terminal carboxylic acid groups. Instead, alkylation of only the phenol was achieved.

In an attempt to make monitoring the reaction by TLC and purification of the reaction mixture by silica chromatography simpler, the corresponding methyl ester was used in place of the carboxylic acid. However, upon repeating the reaction and allowing to heat at reflux for 48h only alkylation of the phenol was observed. A wide range of conditions can be used in the alkylation of phenols and hindered alcohols, such as increasing the strength of the base, the reactivity of the electrophile and the solvent. A summary of the alternative conditions tried are shown in **Table 6**. Unfortunately, all other conditions tried yielded only alkylation of the phenol (**153**).

Initial attempts to yield the dialkylated product involved the addition of tetrabutylammoniumiodide (TBAI) to the reaction mixture (entry 2). If TBAI is present it should convert the alkyl bromide to an alkyl iodide in a Finklestein type reaction. Iodide is a better leaving group than bromide, making the alkyl iodide a better electrophile in the alkylation reaction compared to the alkyl bromide. However, the reaction with TBAI added still yielded monoalkylation of the phenol only (**153**). The solvent was changed to DMF, which has a higher boiling point than THF, and so the reaction could be performed at a higher temperature, however only monoalkylation of the phenol was obtained (entry 3). The use of KH, a stronger base than NaH, still only revealed monoalkylation (entry 4). Finally, the alkyl iodide

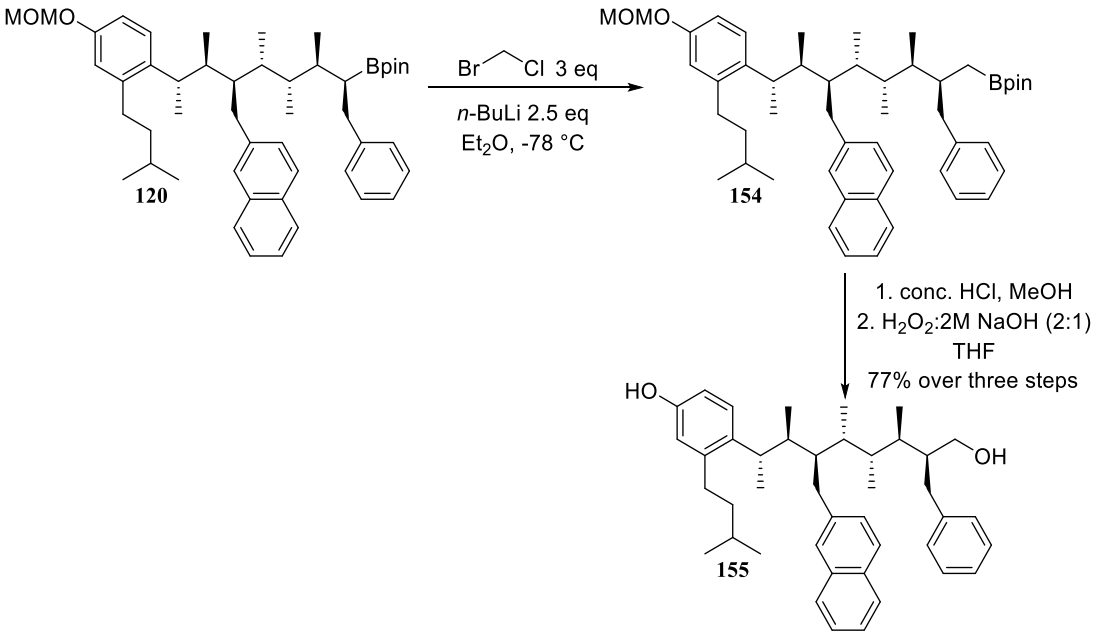
was used directly in the reaction instead of using the alkyl bromide and forming the alkyl iodide *in situ* with TBAI, however the desired reaction was still unsuccessful (entry 5).

Table 6. The conditions used in the attempted alkylation of the bisalcohol **119**



Entry	Reagent	Base	Solvent	Temp (°C)	Additive
1	X=Br	NaH	THF	66	-
2	X=Br	NaH	THF	66	TBAI
3	X=Br	NaH	DMF	120	TBAI
4	X=Br	KH	THF	66	TBAI
5	X=I	NaH	THF	66	-

In an attempt to make the secondary alcohol less sterically hindered and increase reactivity, a Matteson homologation was performed on the final boronic ester **120**. The Matteson homologation adds a methylene group to the carbon chain, and after oxidation of the resulting boronic ester, will yield a primary alcohol.[268] The Matteson homologation typically occurs with >95% conversion and NMR of the crude reaction material revealed full conversion of the starting boronic ester **120** to the homologated boronic ester **154**. The crude reaction mixture was subjected to deprotection of the MOM group and oxidation of the boronic ester, using the conditions described in the section 9.3 of the supporting information, to yield the bisalcohol **155** (**Scheme 19**).

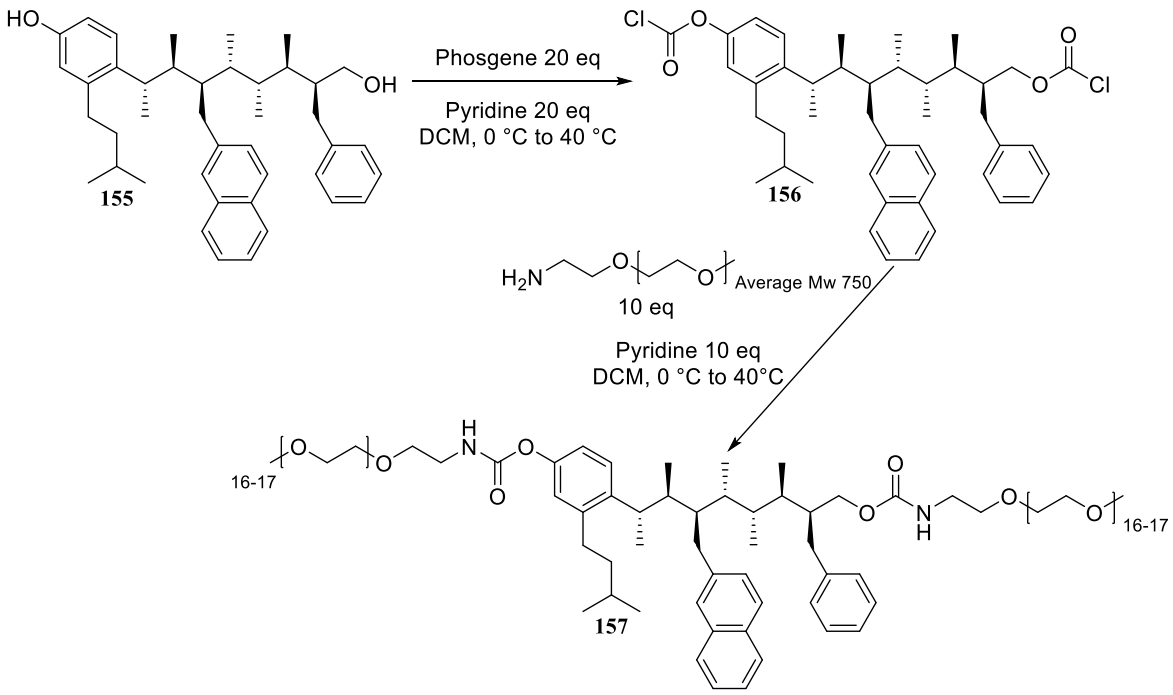


Scheme 19. The Matteson homologation followed by MOM deprotection and oxidation of the boronic ester.

With a primary alcohol now present, the alkylation should hopefully proceed more rapidly than the secondary alcohol. The alkylation of **155** was performed using the conditions listed in entry 1 and entry 2 of **Table 6**. However, both alkylation attempts were unsuccessful and yielded only monoalkylation of the phenol. With no success made in the alkylation of both the phenol and the alcohol to form a bis functionalised product, the possibility of forming the corresponding acyl chlorides from the alcohols was investigated. The acyl chloride is highly electrophilic and should react with a large range of nucleophiles.

A strategy that has proved successful when attempting to improve aqueous solubility of lipophilic ligands is the addition of polyethylene glycol (PEG) chains. In addition to improving aqueous solubility, PEGylated therapeutics typically demonstrate longer circulating half-lives, lower enzymatic degradation, improved resistance from renal clearance, lower peak plasma concentrations and reduced immunogenicity, antigenicity and toxicity.[269–271] Due to the favourable properties that PEG chains can exert on a therapeutic molecule, the use of a PEGylated nucleophile capable of reacting with the acyl chloride was explored. Due to their importance in medicinal chemistry, a large catalogue of functionalised PEG chains are commercially available. Methoxypolyethylene glycol amine (average Mw = 750 $gmol^{-1}$) was chosen as the nucleophile.

Formation of the acyl chloride at both the phenol and the alcohol was first attempted with triphosgene and pyridine in DCM.[272,273] However, the reaction was unsuccessful and incomplete formation of the acyl chloride was observed by TLC. Using phosgene, as a solution in toluene, successfully yielded full conversion of both hydroxy groups to the corresponding acyl chlorides, which was confirmed by TLC, to afford compound **156** (**Scheme 20**). Upon completion of the reaction, the solvent was removed *in vacuo* to remove the excess phosgene, and the crude reaction mixture resuspended in DCM. Pyridine was added to the solution dropwise at 0 °C, followed by a solution of methoxypolyethylene glycol amine in DCM. The reaction was heated to 40 °C and monitored by TLC (**Scheme 20**). Upon completion of the reaction, the crude reaction mixture was filtered through a plug of silica and eluted using 10% MeOH in DCM. The bisPEGylated product **157**, was purified by reverse phase-HPLC in 11% yield. With a bisfunctionalised molecule in hand that should display improved water solubility the binding of this mimetic to Mdm2 by common biophysical methods can be explored.

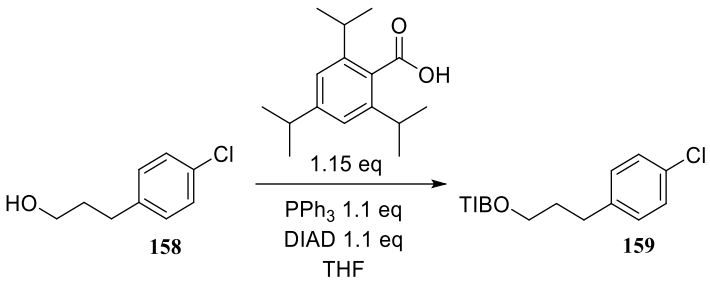


Scheme 20. Improving the water solubility by incorporating PEG chains to form the bisPEGylated product **157**

4.3 The Synthesis of a Library of p53 Mimetics

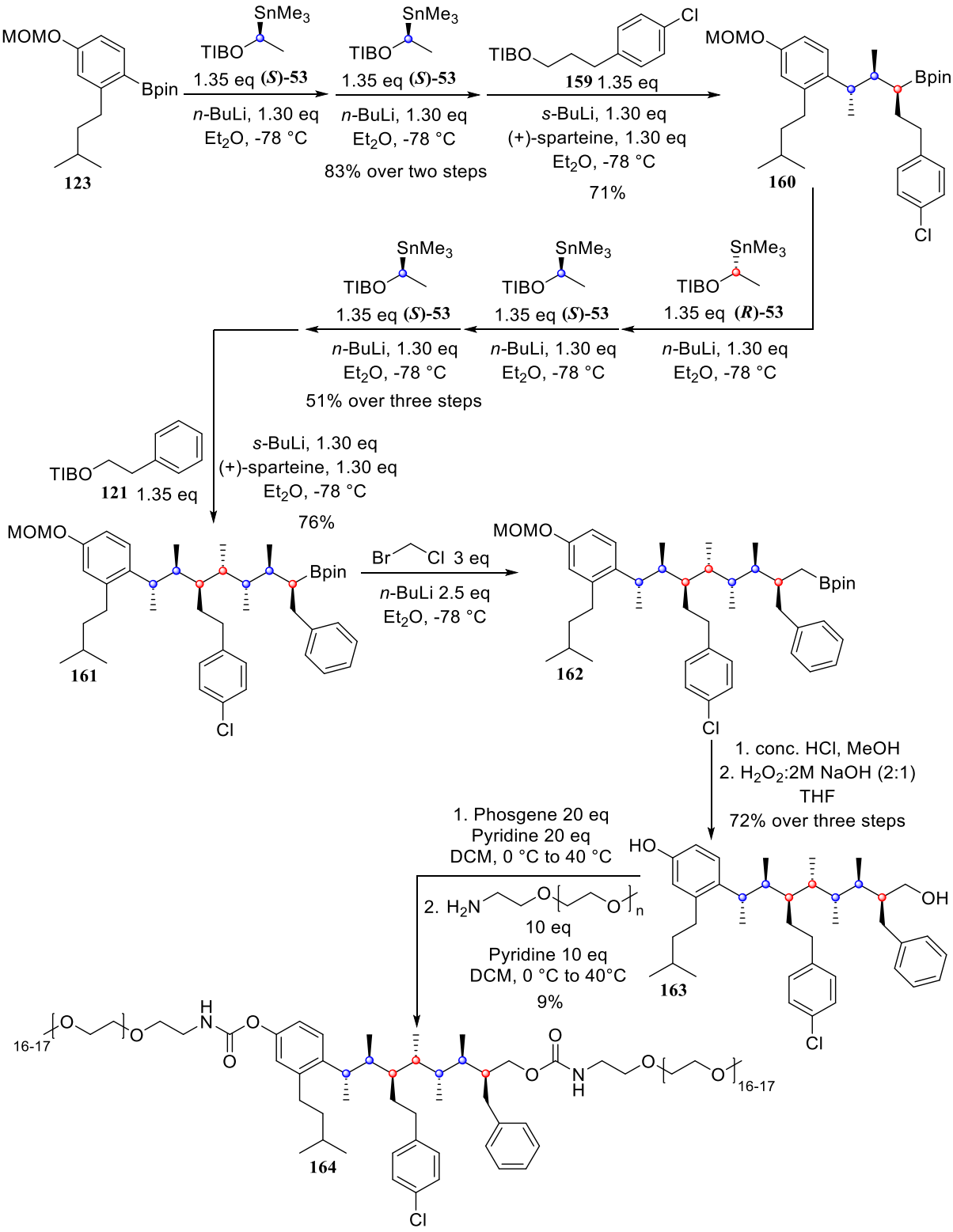
Chapter 3, demonstrated that by using molecular docking the binding affinity of the designed p53 mimetic **96** to Mdm2 could be increased by substituting the naphthyl group for a *para*-chloro substituted phenyl group (**105**)**.** The positive effect of the chlorine atom on the binding affinity has also been reported by many others.[196] It was also demonstrated that the introduction of fluorine atoms onto the scaffold (**106**) increased the binding affinity further. Using lithiation–borylation, the synthesis of a mimetic with a different group in place of the naphthyl group is accomplished simply by substituting a different benzoate ester in the sequence. This will allow for a library of p53 mimetics to be synthesised using one synthetic methodology.

The second member of the library of p53 mimetics to be synthesised was based on **105**, featuring a *para*-chloro phenyl group in place of the naphthyl group. This can be achieved by replacing the naphthyl benzoate ester (**122**) used in the synthesis of **157** with a benzoate ester featuring an appropriate side chain. The synthesis of the benzoate ester used in the sequence (**159**) is shown in **Scheme 21**.



Scheme 21. The synthesis of the benzoate ester **159** used in the synthetic sequence of a second p53 mimetic

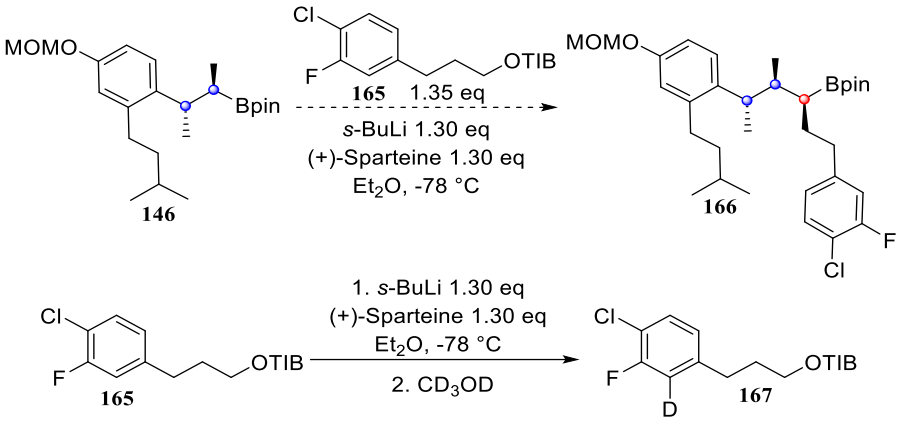
A similar lithiation–borylation sequence was performed as described previously for **120**. However, in addition to the first and second being performed iteratively, for the synthesis of **161**, the fourth, fifth and sixth homologations were also performed iteratively. Following the introduction of the hot-spot groups and the methyl groups, a Matteson homologation was performed as described previously, to install the methylene group, affording compound **162**. This was followed by the deprotection of the MOM group and oxidation of the boronic ester to yield the bisalcohol **163**. The bisalcohol **163 was** converted to the corresponding bisacyl chloride with phosgene, as described previously. Subsequent reaction with methoxypolyethylene glycol amine gave the form the bisPEGylated derivative **164** (**Scheme 22**). The PEG derivative **164** was purified by reverse phase-HPLC in 9% yield and concentrated ready for biological testing.



Scheme 22. The synthesis of a second p53 mimetic **164** using iterative lithiation–borylation.

Following the successful synthesis of **164**, we turned our attention to substituting the *para*-chloro phenyl ring with a fluorine atom, resembling the designed p53 mimetic **106**. Unfortunately, the attempted lithiation of the benzoate ester **165** did not yield lithiation at the desired α-position and as a result the homologation with boronic ester **146** was unsuccessful. Quenching the lithiation reaction with

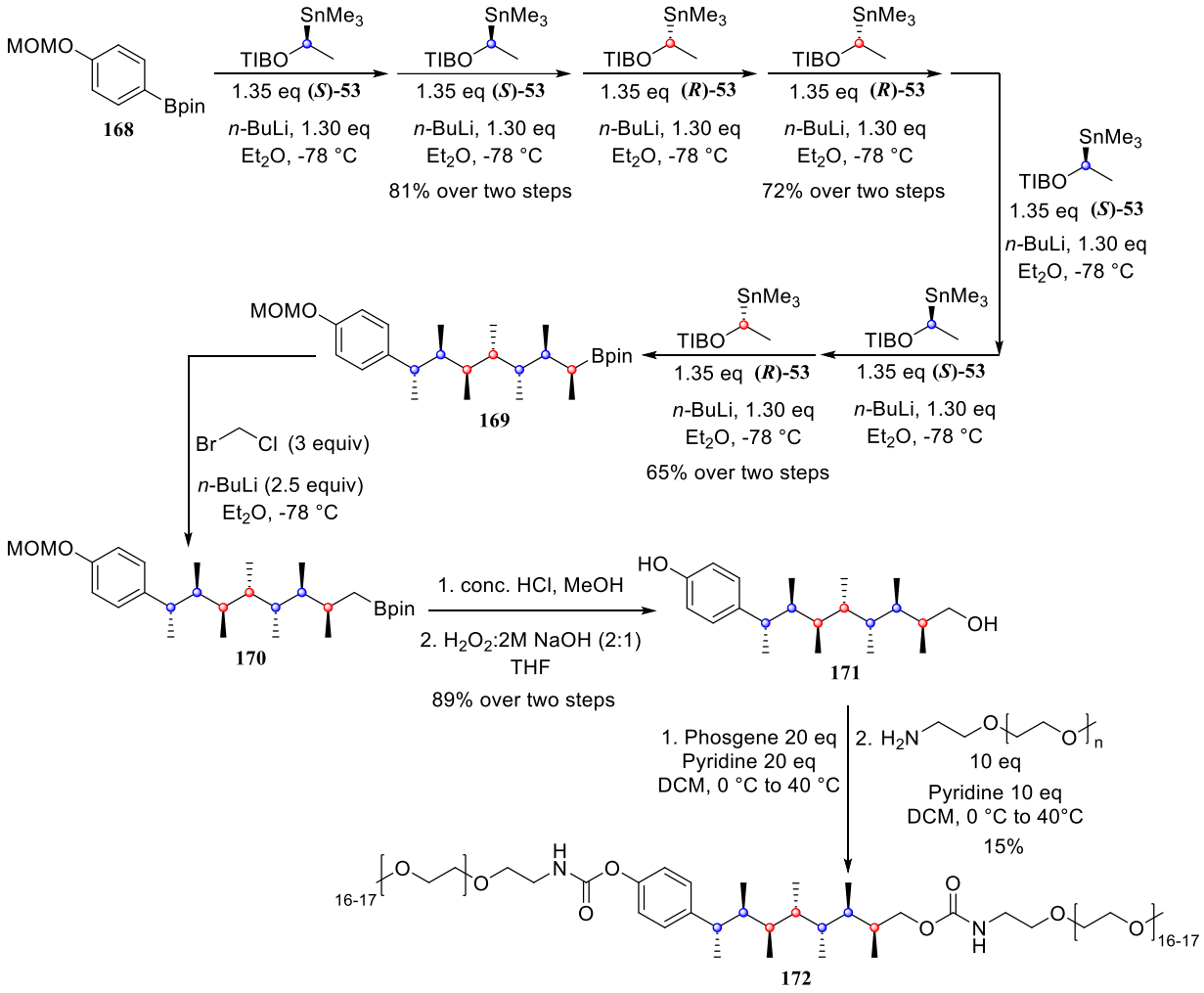
CD_3OD revealed lithiation at the aromatic proton adjacent to fluorine, confirmed by full deuteration of this position observed by NMR. This suggests that the fluorine directs lithiation to the *ortho* position for benzoate ester **165** (**Scheme 23**). Strategies to overcome this, such as blocking the proton α- to fluorine or masking the fluorine itself and revealing it after the lithiation–borylation sequence will be explored in future work.



Scheme 23. The unexpected lithiation of benzoate ester **165**

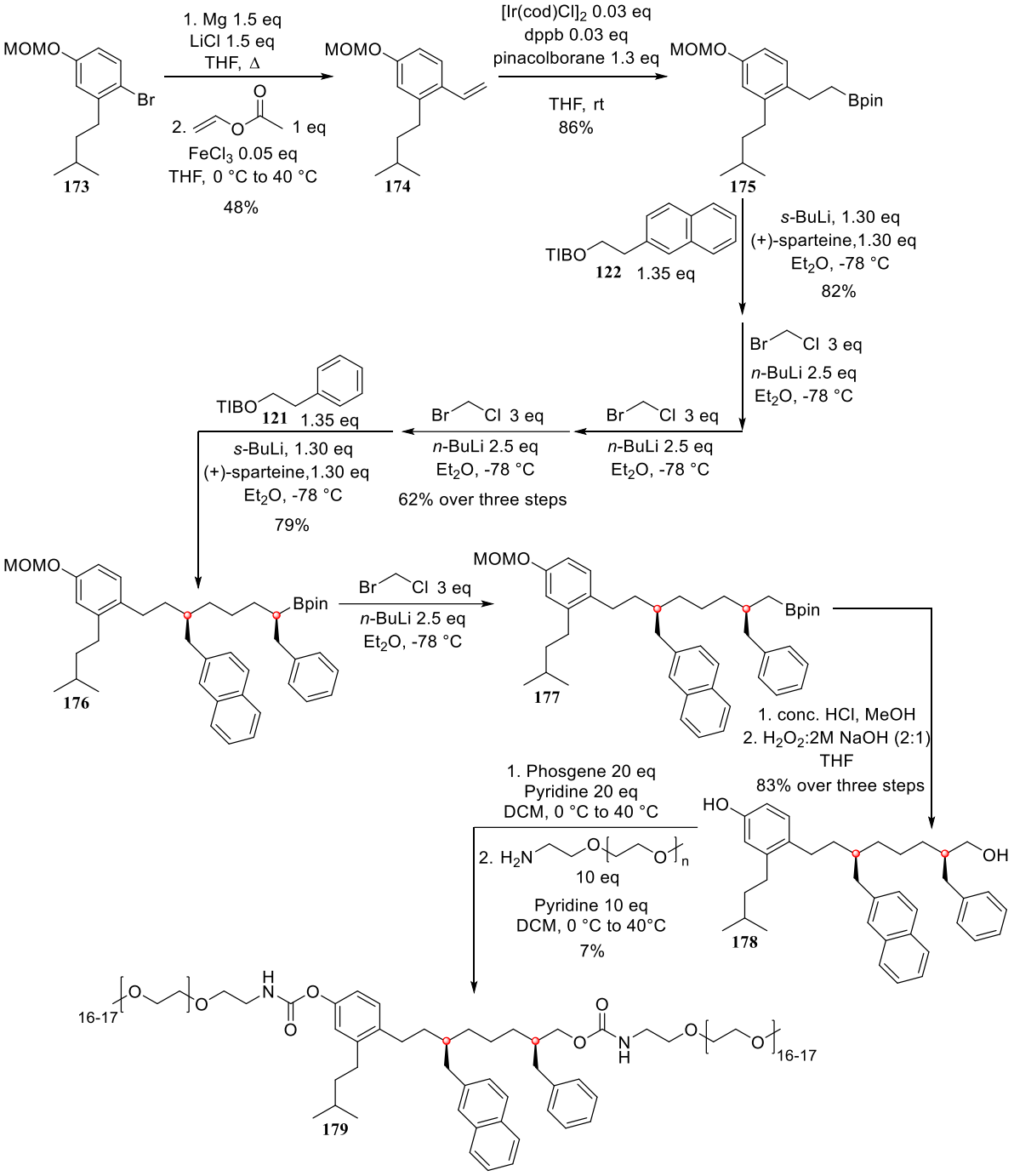
4.4 The Synthesis of Control Molecules

Chapter 3 also discussed the need for two control molecules, to confirm the importance of conformational control and the hot-spots in obtaining higher affinity therapeutics. The two control molecules can also be synthesised by iterative lithiation–borylation. The first control molecule synthesised was **172**, which featured no hot-spot groups. This molecule was synthesised in a very similar fashion to that reported by Aggarwal *et. al.* The synthesis of control molecule **172** is shown in **Scheme 24**. Although the molecules synthesised by Aggarawal *et. al.* were performed fully iteratively with no column chromatography, for the control molecule **172**, column chromatography was performed after the second, fourth and seventh homologation to ensure no under-homologated product was carried through the reaction. Upon completion of the synthesis of the final boronic ester **170**, the phenol was deprotected and the alcohol was oxidised as performed previously for the two p53 mimetics. The resulting bisalcohol, **171** was subjected to reaction with phosgene to form the acyl chloride, followed by reaction with methoxypolyethylene glycol amine to form the bisPEGylated derivative **172**. The PEG derivative **172** was purified by reverse phase-HPLC in 15% yield, and concentrated ready for biological testing.



Scheme 24. The synthesis of control molecule **172** using lithiation–borylation

The second control molecule, **179**, features no methyl groups and thus should display no conformational control. This synthesis of the control molecule is shown in **Scheme 25**. The first step of the synthesis involves an iron-catalysed cross coupling reaction between an alkenyl acetate and the aromatic bromide species **173** to form the trisubstituted phenyl **174** in 48% yield. An iridium catalysed hydroboration is used to install the pinacol boronic ester, to give the starting boronic ester **175** in 86% yield. The boronic ester **175** can then undergo a lithiation–borylation sequence, as shown in **Scheme 25**. Matteson homologations are interspersed with homologations to introduce the two hot-spot groups using sparteine-mediated deprotonation of the corresponding benzoate esters, as described previously. The final boronic ester **177** was subjected to the MOM deprotection, oxidation of the boronic ester and PEGylation as seen for the previous p53 mimetics the other control molecule. The PEG derivative **179** was purified by reverse phase-HPLC in 7% yield and concentrated ready for biological testing.



Scheme 25. The synthesis of control molecule **179**

With the synthesis of two p53 mimetics (**157** and **164**) and two control molecules (**172** and **179**) bearing groups that should improve their water solubility, their binding to Mdm2 can be explored. Prior to this however, the conformational bias of these molecules in solution will be confirmed using a hybrid NMR and QM calculations method. This will be discussed in the next section.

Chapter 5: Using NMR Spectroscopy and Computation to Analyse the Conformation of the Designed α-Helix Mimetic of p53

Section 3.2 of Chapter 3, discussed the use of MM to explore the conformational landscape of a suitable mimetic of p53. MM revealed a linear conformational bias for the designed ligand **96** with over 74% of conformers adopting a linear conformation. This conformational bias can be confirmed experimentally using a hybrid NMR spectroscopy and computational approach. This approach involves comparing experimental NMR parameters, including both ^{1}H-^{1}H and ^{1}H-^{13}C scalar coupling constants, interproton distances and chemical shifts, to those that are calculated using DFT calculations.[109] Following an initial stochastic MM conformational search, a redundant conformer elimination is performed which removes conformers that are identical in energy and whose cartesian coordinates differ by less than a predefined distance. The resulting conformers are taken through to QM calculations using DFT, which calculates the Gibbs free energies and resulting Boltzmann populations of each conformer, along with their GIAO-based NMR parameters. The NMR parameters are Boltzmann averaged using the calculated conformer populations, to provide the ensemble-averaged NMR properties of the molecule, which can be compared to the experimental NMR data. The workflow employed in this project for this procedure is shown in **Figure 66.**

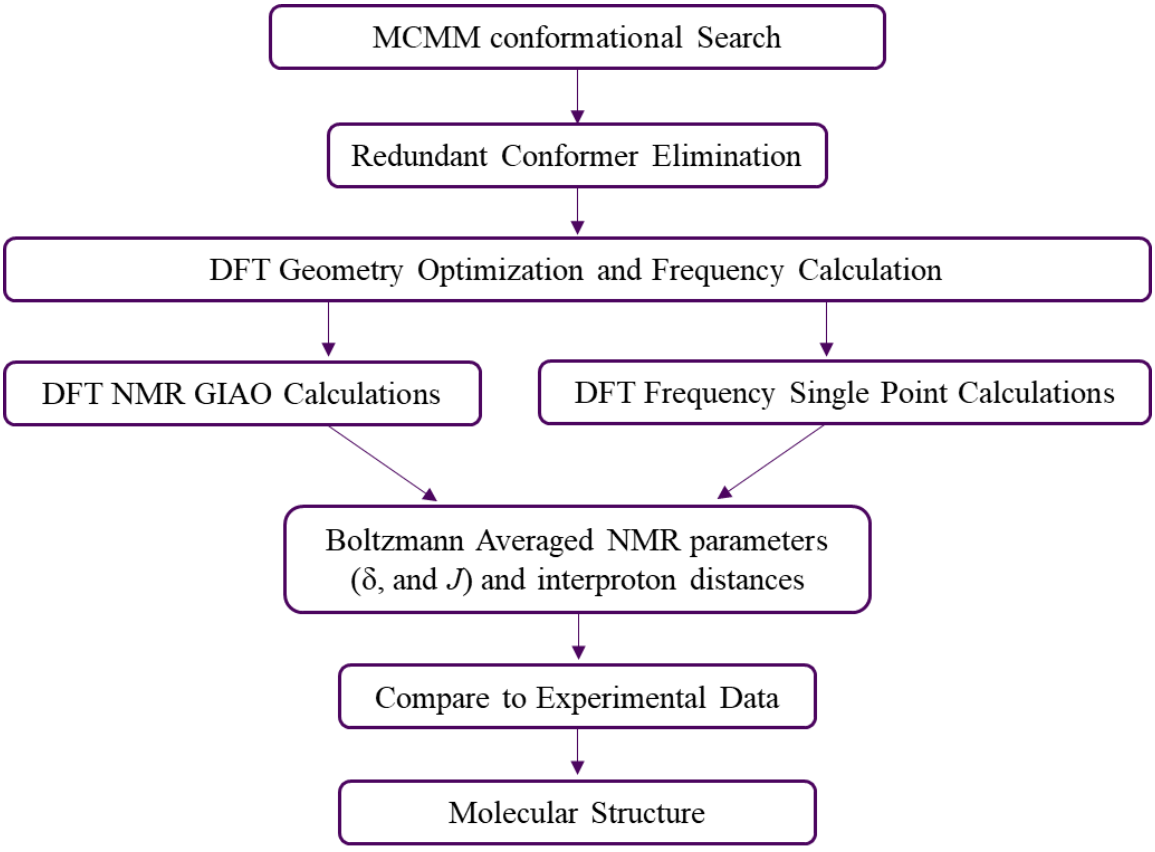


Figure 66. The computational workflow employed in this project.

5.1 The Choice of a Representative Structure

Although five ligands have been synthesised (**119, 155, 157, 163, 164**, **Figure 67**), their structural similarity and computational demand of the workflow described in **Figure 66**, led us to validate the structure of them all by analysis of the conformation of a single representative structure. The full DFT-based conformational analysis will therefore be performed on a representative structure however, the conformation of the other structures will be considered from examination of their NMR spectra and comparison to MM conformational searching.

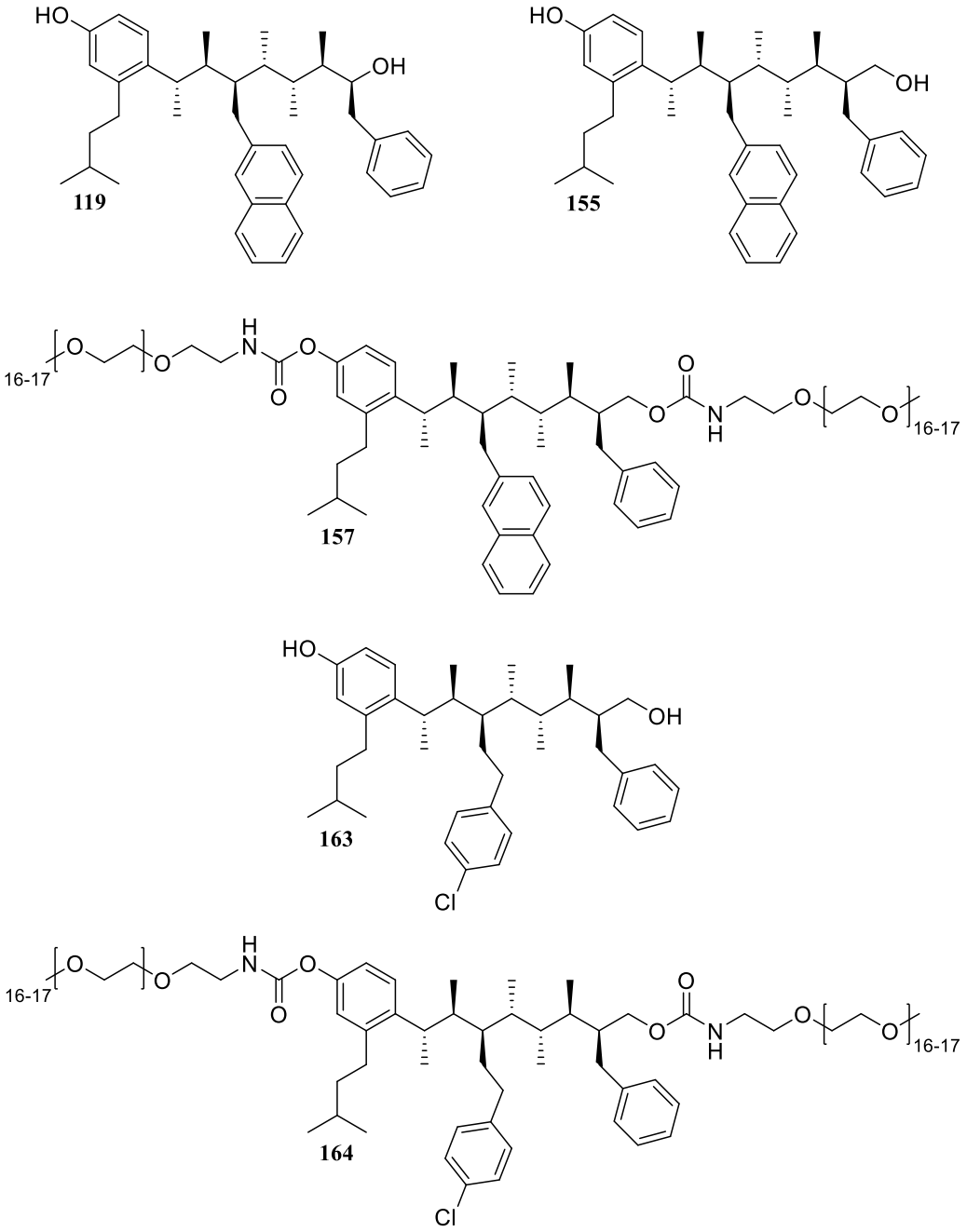


Figure 67. The important ligands synthesised previously, one of which will be chosen to perform a full conformational analysis.

An ideal candidate for the conformational analysis will have a well dispersed NMR spectrum with little or no overlap. This will aid with assignment, measuring the scalar coupling constants and obtaining interproton distances from selective 1D-NOE spectroscopy. The PEG derivatives **157** and **164** are not ideal candidates due to the signals from the PEG chains dominating the NMR spectrum. Additionally, the peaks are very broad due to fast T_2 relaxation (relaxation in the *x*-*y* plane). T_2 relaxation, known as spin-spin relaxation, is related to line with by $\nu_{1/2} = 1/\pi T_2$ where $\nu_{1/2}$ = width at half height. Thus, fast relaxation in the *x*-*y* plane leads to broader peaks. The value of T_2 decreases as molecular size or solvent viscosity decreases since it is related to the rotational correlation time, τ_c, which time describes the average time taken for the molecule to rotate through one radian. Therefore, the PEG derivatives are expected to have the shortest T_2 and thus broadest peaks. Furthermore, it is likely that the tumbling of the molecules is anisotropic in solution, i.e. different rotational axes of the molecule will have different values of τ_c. This anisotropic motion also reduces T_2 leading to broad peaks.

An additional negative effect of the high molecular weight of the PEG derivatives, which have a molecular weight of >2000 $gmol^{-1}$, is that the NOE intensity approaches zero and becomes negative as molecular weight increases. (**Figure 68**) As molecular weight increases, τ_c increases and molecular motion becomes slower than the Larmor frequency. In cases such as this, the zero quantum (ω_0) cross relaxation pathway dominates and a negative NOE results. The measured NOEs of the PEG derivatives will likely not be at a maximum negative NOE, which is not ideal for the accurate quantitative analysis of interproton distances using NOE. Additionally, the signals of the CH backbone peaks for the PEG derivatives are not well dispersed.

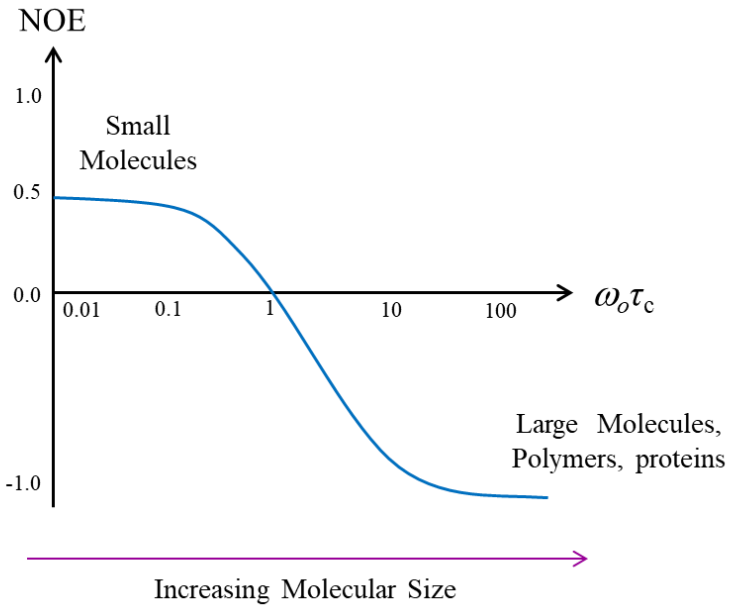


Figure 68. The effect of increasing molecular weight on the NOE intensity.

The NMR spectra of the three remaining candidates, **119**, **155**, and **163**, were examined to identify the best structure for the conformational analysis. The spectral dispersity of the backbone CH protons along the is the most important parameter when considering the choice of a representative structure for several reasons. The first reason for this is that it makes it easier to measure the $^{3}J_{HH}$ coupling constants along the backbone which provide valuable conformational information along the backbone, due to the relationship between $^{3}J_{HH}$ values and dihedral angle, as described by the Karplus relationship.[87,88] Furthermore, a well resolved spectra is important for selective irradiation in 1D-NOE experiments used to determine accurate interproton distances. In selective 1D-NOE methods, a peak generally needs to be separated from other peaks in the spectrum by ~50 Hz to ensure selective irradiation. The NMR spectra of the naphthyl derivative **119** was the most dispersed spectra with each CH proton along the backbone giving a separated, or partially separated, resonance (**Figure 69**). The NMR spectra of the Matteson homologated naphthyl derivative **155** and the para-Cl phenyl derivative **163** revealed considerable overlap between the backbone CH protons and were therefore not ideal for use in the full DFT-based conformational analysis. Thus, our analysis will be performed on the naphthyl derivative **119**.

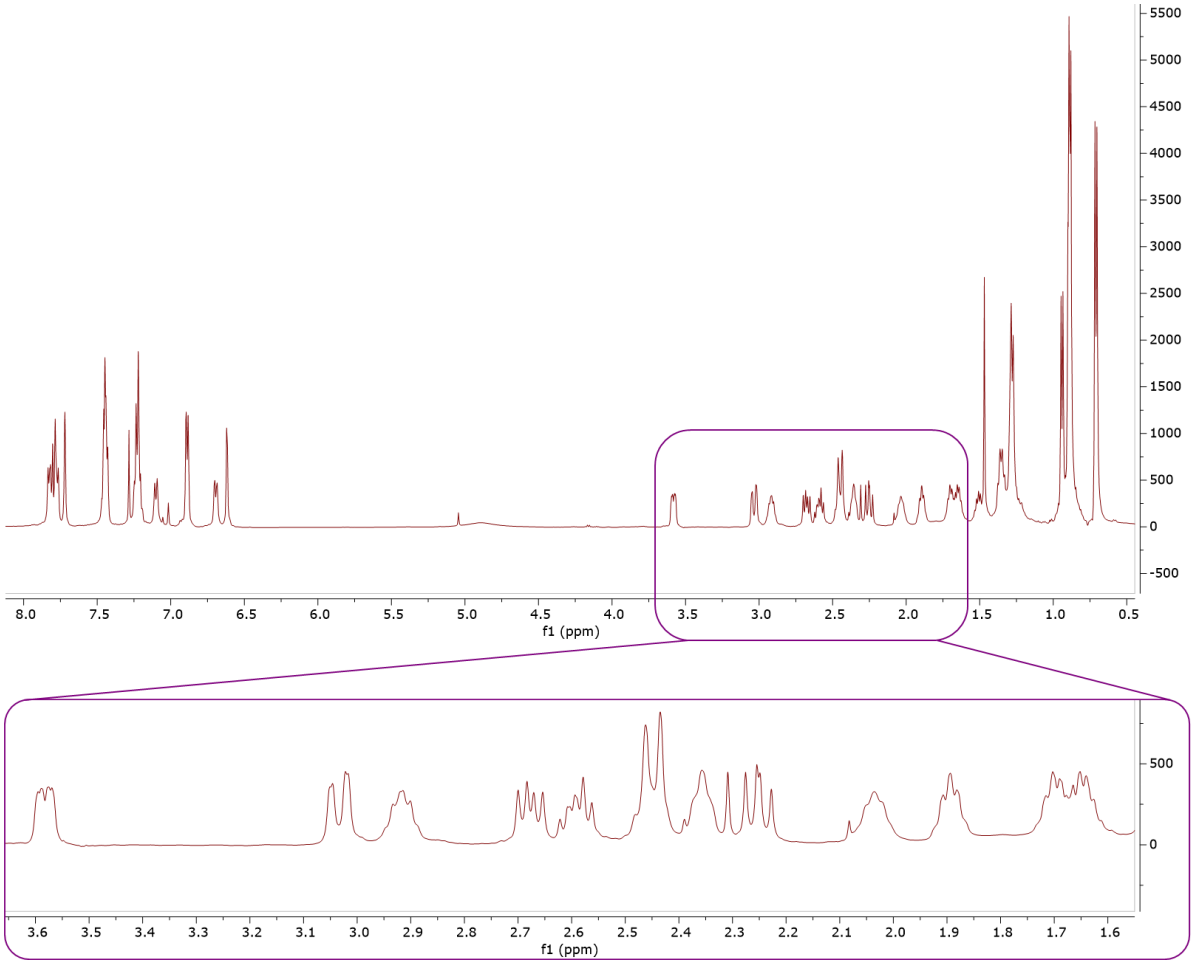


Figure 69. The ^{1}H NMR spectra of the naphthyl mimetic **119**, with an expansion of the backbone CH region showing well dispersed peaks.

5.2 The Calculation of Conformer Geometries, Energies and NMR Parameters using DFT Calculations

Section 3.2 of Chapter 3 discussed the use of MM to explore the conformational landscape of a possible p53 mimetic and found that structure **96** displayed a linear conformational bias with over 74% of conformations adopting a linear conformation. Although MM is a useful tool in identifying the overall conformational bias of a molecule, it is not adequate at accurately predicting conformer energies and corresponding Boltzmann populations. Therefore, DFT will be used to further optimise conformer geometries, and calculate conformer energies and NMR parameters.

The conformational search of **119** (**Figure 70**) was performed in chloroform using the MMFFs forcefield and MCMM searching method. A total of 17380 conformers were found within 21 kJ mol^{-1} of the global minimum. The MM results revealed that for **119**, 77.5% of the conformers adopted a linear conformation. DFT calculations cannot be performed on all 17380 conformers, as this is too demanding for the computational resources available. The number of conformers was first reduced by performing a redundant conformer elimination, during which redundant conformers are eliminated using maximum atom deviations. If the sum deviation in the positions of the selected atoms differs by more than 0.5 Å it will be retained as a unique conformer. Only atoms in the backbone and first carbon of the side chains were selected and as a result, rotation of the side chains past the first carbon atom will be poorly sampled. Although this step results in parts of the molecule being poorly described, the backbone of the molecule will still be well explored. The redundant conformer elimination reduced the number of conformers within 30 kJ mol^{-1} to 1115.

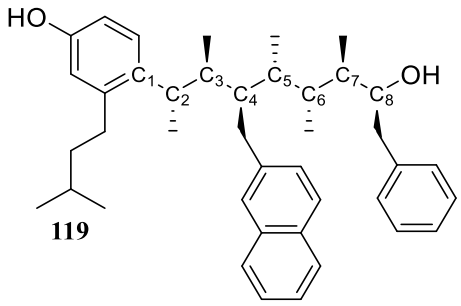


Figure 70. The structure of ligand **119**, with the dihedral angles of the backbone annotated.

The relative potential energies obtained from MM for the 1115 conformers were used to calculate the Boltzmann population of each conformer. The results are summarised as a bubble plot, shown in **Figure 71**). The clear bias towards dihedral angles of 180° along the backbone is indicative of a linear conformational bias. The lowest energy conformer has a population of 22% and displays a linear conformation.

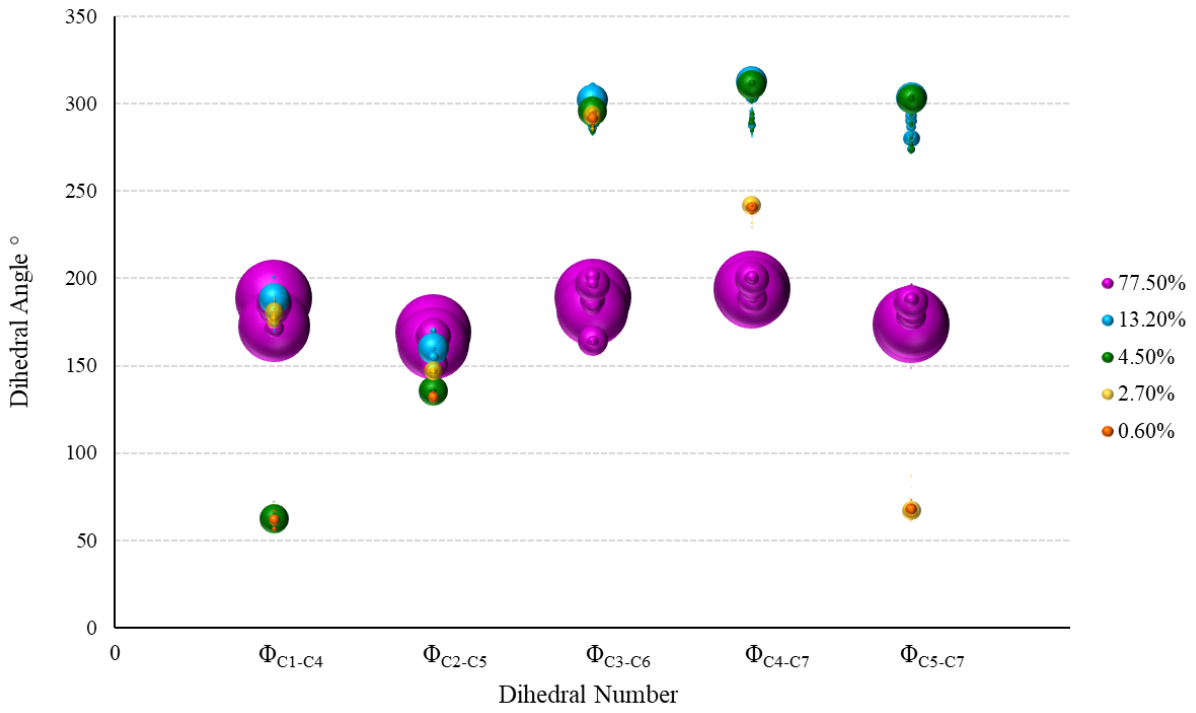


Figure 71. The results of the MM conformational search of ligand **119**, summarised as a bubble plot. Dihedral angles of 180° along the backbone are clearly favoured.

Because of the time demands of DFT computation (hours to days per conformer) compared to MM (second to minutes per conformer), a further reduction in the number of conformers was required prior to DFT calculations, so only conformers that comprised the top 99.9% of Boltzmann populations (241 conformers) were selected. All 241 conformers were subjected to DFT geometry optimisation and frequency calculation using mPW1PW91/6-31g (d) basis set (basis set 1). All DFT calculations were performed in chloroform using the implicit solvent model, IEFPCM. The functional was chosen as it has been shown by Bifulco *et al.* to be the most accurate functional for computing chemical shifts and $^{n}J_{CH}$ scalar coupling constants.[274] Although a triple ζ basis set has been shown, by Dr. Zhong at the University of Bristol, to provide the greatest accuracy when Boltzmann averaged NMR parameters were compared to calculated values, it comes at significant computational cost.[275] Therefore, a double ζ basis set was used for the first geometry optimisation and frequency calculation and a triple ζ basis set for the subsequent single point energy calculation, which will provide a more accurate estimation of the Gibbs free energy and also correct for the zero point energy. Additionally, polarisation wavefunctions (d,p) were included to allow for charge polarization in the single point calculation. Again, adding the polarisation functions adds significant computational cost, so the effect of adding these to the calculation is estimated using a single point energy calculation. Of the 241 conformers subjected to the first DFT calculation, 9 conformers did not converge during the geometry optimisation and were eliminated. Additionally, 5 conformers were found to possess an imaginary frequency and were eliminated from the conformer pool. An imaginary frequency occurs if the conformer geometry exists

in a saddle point and not a local minimum on the potential energy surface and thus is not a stable structure and can be removed.

Following the first geometry optimisation and frequency calculation, the 227 converged conformers were subjected to a single point energy calculation using mPW1PW91/6-311g (d,p) (basis set 2). The Gibbs free energies of each conformer were estimated using **Equation 6**.

$$G_i^{Basis\ Set2} = E_i^{Basis\ Set2} + (G_i^{Basis\ Set1} - E_i^{Basis\ Set1})$$

Equation 6. The estimation of the Gibbs energy of conformer *i*.

Where $G_i^{Basis\ Set2}$ is the estimated Gibbs energy of conformer *i* using the 6-311g(d,p) basis set, $E_i^{Basis\ Set2}$ is the potential energy of conformer *i* calculated using the 6-311g(d,p) basis set, $G_i^{Basis\ Set1}$ is the calculated Gibbs energy of conformer *i* using the 6-31g(d) basis set and $E_i^{Basis\ Set1}$ is the potential energy of conformer *i* calculated using the 6-31g(d) basis set.

The calculated Gibbs energies were used to recalculate the Boltzmann distribution of the conformers for ligand **119**. A slight increase in the percentage of conformers adopting a linear conformation from 77.5% to 80.8% was observed after going from MM to QM (**Figure 72**).

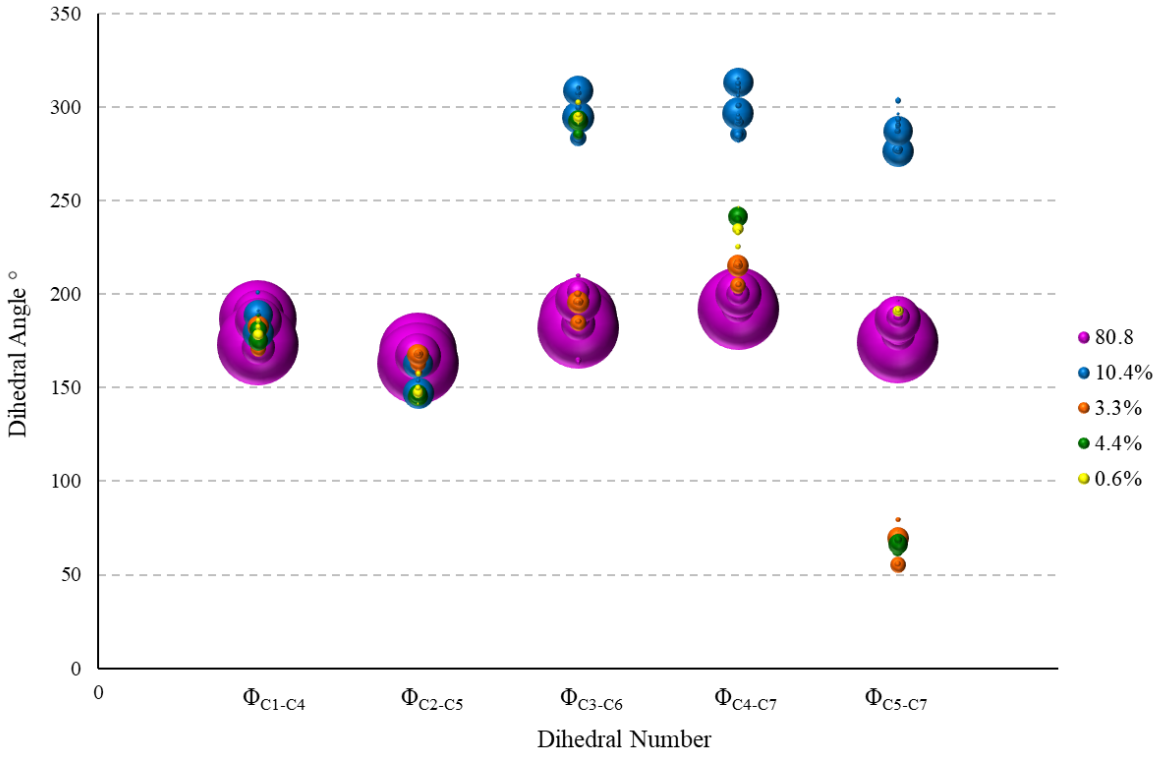


Figure 72. A summary of the Boltzmann distribution of conformers after DFT calculations for ligand **119**.

Following the calculation of conformer energies and Boltzmann populations, the calculation of Boltzmann averaged calculated NMR parameters and interproton distances was performed. The interproton distances were extracted from the optimised geometries and were Boltzmann averaged against their calculated populations to provide the ensembled averaged interproton distances. These distances can then be compared to those derived using 1D-NOE spectroscopy to provide an indication towards the conformational landscape of **119** in solution. The quantitative interproton distance analysis by 1D-NOE is highly sensitive to conformer populations since interproton distances scale with r^{-6} and thus it is very important that the calculated Boltzmann populations are accurate and that all conformers subjected to DFT are included in the extraction of the calculated ensembled averaged interproton distances. Therefore all 227 converged conformers were used in the interproton distance analysis, discussed in section 5.4.

The final stage of the DFT procedure is to calculate the NMR parameters. Magnetic shielding tensors and spin-spin coupling constants were calculated with the GIAO method using mPW1PW91/6-311g (d,p). Decontracted core orbitals were used, called with the Gaussian keyword 'mixed', to improve the description of electron density at the nucleus by mixing the core orbitals into the Fermi contact term. This has been shown to improve the accuracy of calculated scalar couplings.[276] The calculation of NMR parameters is computationally extremely intensive and as a result a reduced subset of conformers were subjected to the NMR calculations. Scalar coupling constants and chemical shifts are not nearly as sensitive to conformer populations compared to interproton distances from 1D-NOE, and thus this reduction in conformers is justified. The Boltzmann distribution after DFT was used to select conformers for the NMR calculations. The top 99% of conformers after DFT were taken through to the NMR calculations, corresponding to 81 conformations. The magnetic shielding tensors and scalar coupling constants were Boltzmann averaged against their Boltzmann populations. The magnetic shielding tensors were converted to a chemical shift by referencing the magnetic shielding tensors to those calculated for tetramethylsilane (TMS) at the same level of theory (**Equation 7**).[135]

$$\delta_{calc}^{x} = \frac{\sigma^{O} - \sigma^{x}}{1 - \sigma^{x}/10^{6}}$$

Equation 7. The calculation of chemical shifts from magnetic shielding tensors.

Where δ_{calc}^{x} is the calculated chemical shift for nucleus *x* (in ppm), σ^{x} is the Boltzmann averaged magnetic shielding tensor for nucleus *x* and σ^{O} is the magnetic shielding tensor for TMS.

The calculated Boltzmann averaged interproton distances, scalar coupling constants and chemical shifts can be compared to those derived experimentally by NMR spectroscopy. If the conformational landscape of **119** has been adequately described by DFT calculations, a good fit between experimental and calculated data should be observed, suggesting that a linear conformation is preferred in solution.

If, however, the experimental and calculated data does not correlate well, it suggests that the Boltzmann populations of the conformers has not been correctly estimated by DFT and the conformational landscape of **119** may not be what computation suggests.

5.3 A Comparison of the Experimental ^{1}H-^{1}H and ^{1}H-^{13}C Scalar Coupling Constants to the Boltzmann Averaged Calculated Values

As discussed in the introduction, both the ^{1}H-^{1}H and ^{1}H-^{13}C scalar coupling constants provide valuable information regarding the conformation of a molecule, primarily due to the relationship of vicinal $^{3}J_{XH}$ (where X = H or C) with dihedral angle, known as the Karplus relationship. The conformation of **119** was first explored by examining the $^{3}J_{HH}$ coupling constants along the backbone of the molecule. If a linear conformation is favoured the HCCH dihedral angles along the backbone of the molecule should alternate between 180° and 60° corresponding to an alternating *anti* - *gauche* relationship along the backbone. Thus, alternatingly large and small values of $^{3}J_{HH}$ should be observed based on the Karplus relationship **(Figure 73)**.

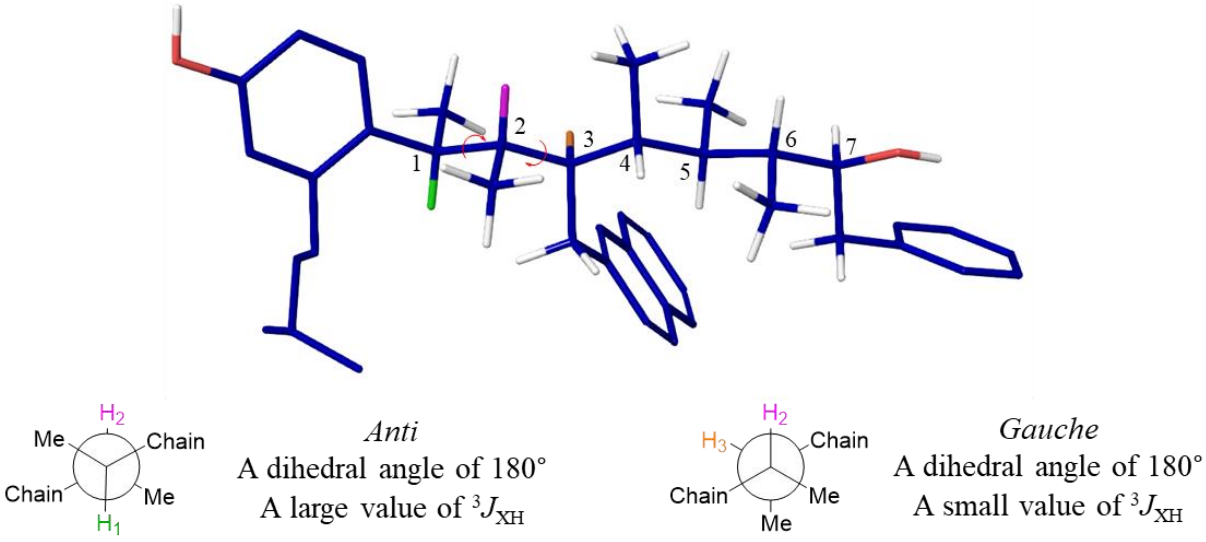


Figure 73. The lowest energy conformer of **119**, with the backbone CHs of the first two dihedral angles highlighted. The two Newman projections show that the dihedral angle between H1 (green) and H2 (pink) is 180° and between H2 (pink) and H3 (orange) is 60°. This 180° – 60° – 180° – 60° pattern of HCCH dihedral angles continues along the chain.

Although the CH peaks corresponding to the protons along the backbone ($H_{1\text{-}7}$) are well dispersed, they are very broad and the multiplicities of the peaks cannot easily be distinguished, making the values of $^{3}J_{HH}$ difficult to measure experimentally. Peak broadening can be caused by a number of issues including instrumental problems such as poor tuning and shimming or issues with the sample itself such as sample inhomogeneity or paramagnetic impurities. However, it is more likely that the broad peaks seen in **119** are caused by short T_2 relaxation, as discussed previously.

Despite broadening of the peaks, values of $^{3}J_{HH}$ were measured directly from the 1D ^{1}H NMR spectrum using SpinSimulation; software from MestreNova. During a spin simulation, the expected coupling constant and the exact chemical shift, typically measured from pureshift ^{1}H spectra, are used to simulate the spectra. The simulated spectra are then compared to the measured spectra and the estimated coupling constant altered until the simulated spectrum matches the measured spectrum. Using SpinSimulation, it is very easy to quickly identify whether the value of $^{3}J_{HH}$ is large (> 9 Hz), small (<4 Hz) or an average (=7 Hz) as the rough peak shape will only match to the correct answer. For example in **Figure 74** it is very clear that a small value of $^{3}J_{HH}$ is contributing to the multiplicity of this peak as if a large or average value of $^{3}J_{HH}$ is used in the simulation then the simulated spectrum does not match the experimental spectrum. From this, the coupling constant can be changed by very small increments (0.1 Hz) to find the best fit between simulation and experimental.

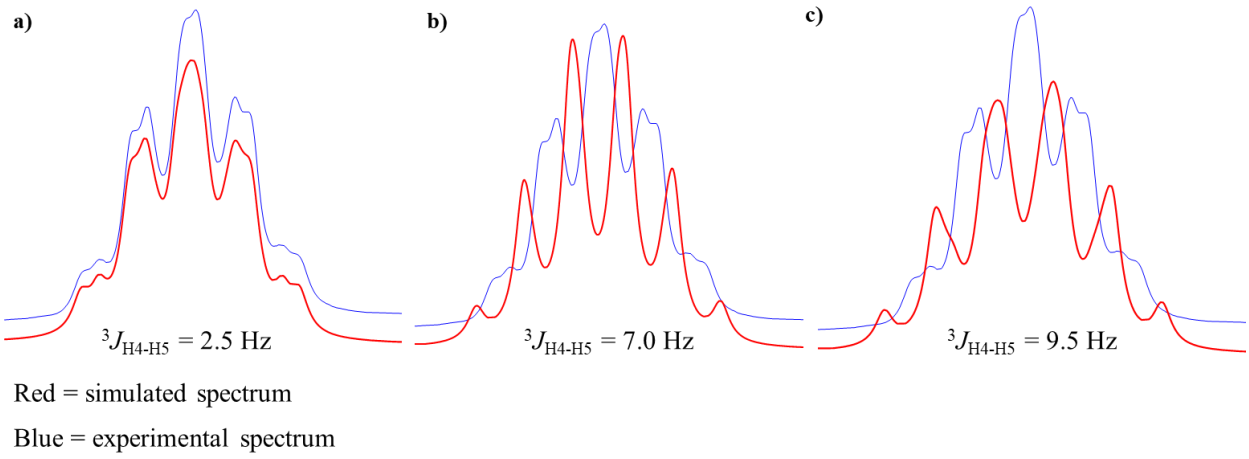


Figure 74. The simulation of the H_4-H_5 peak for **119** with three different values of $^{3}J_{HH}$.

This process was performed for **119**, and a good correlation between the simulated spectrum and the experimental spectrum was obtained. An example of two peaks simulated for **119** is shown in **Figure 75**.

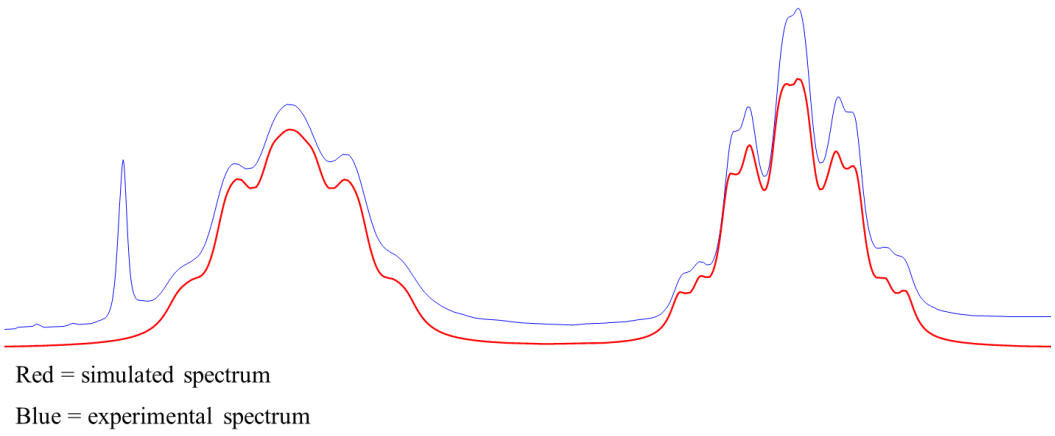
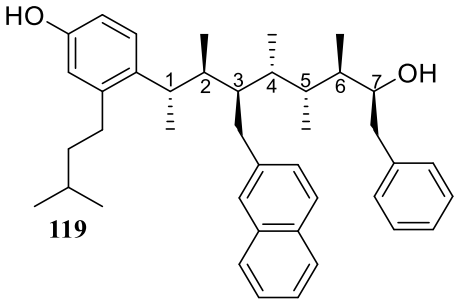


Figure 75. The experimental (blue) and simulated spectrum (red) of two peaks present in the ^{1}H spectrum of **119**.

The experimentally measured values of $^{3}J_{HH}$, obtained from spin simulation, are shown in **Table 7**. It can clearly be seen that the HCCH dihedral angles along the backbone of **119** alternate between large and small, confirming that an *anti-gauche-anti-gauche* relationship exists and suggesting a linear conformation is preferred. This preference can be confirmed by comparing the values obtained experimentally to those calculated by DFT calculations. As discussed in section 1.3.2 of Chater 1, the Boltzmann averaged NMR parameters can be calculated using DFT calculations and can provide a more detailed conformational analysis. A good correlation between experimental and calculated values of $^{3}J_{HH}$ was observed, with a mean absolute deviation (MAD) of 0.48 Hz and a standard deviation (SD) of 0.33 Hz (**Table 7**). This confirms that the populations of the conformers estimated by DFT is in good agreement with the populations of the conformers that exist experimentally at room temperature, confirming the linear conformational bias. As well as considering 3-bond H-H scalar coupling, 2-bond and 4-bond H-H scalar couplings were also measured experimentally and compared to the values calculated by DFT. Comparing all experimental values of $^{n}J_{HH}$ to the Boltzmann averaged values calculated using DFT calculations a MAD of 0.71 Hz and SD of 0.58 Hz was obtained for $^{n}J_{HH}$.

Table 7. The experimental $^{3}J_{HH}$ values measured for **119** using spin simulation. Only the $^{3}J_{HH}$ values along the backbone are shown.



119

H	H	Relationship	Dihedral Angle (°)	Anticipated Coupling (Hz)	Experimental Coupling (Hz)	Calculated Coupling (Hz)*	Absolute Deviation (Hz)
1	2	*Anti*	180	Large >8	10.3	10.7	0.4
2	3	*Gauche*	60	Small <5	2.0	2.6	0.6
3	4	*Anti*	180	Large >8	8.2	9.3	1.1
4	5	*Gauche*	60	Small <5	2.9	3.2	0.3
5	6	*Anti*	180	Large >8	9.2	9.4	0.2
6	7	*Gauche*	60	Small <5	4.6	4.9	0.3
						Mean Absolute Deviation (Hz)	**0.48**
						Standard Deviation (Hz)	**0.33**

*mPW1PW91/6-311g(d,p)

The MAD and SD obtained can be compared to those reported in the literature for similar molecules, to confirm that validity of the analysis. For a rigid molecule a MAD of <0.5 Hz can be expected, however for molecules of increasing flexibility, such as the molecules **10**, **11**, reported by Aggarwal *et. al.*[27] a MAD of ~1 Hz is expected. Therefore, the MAD and SD obtained in the comparison of experimental values of $^{n}J_{HH}$ to Boltzmann averaged calculated values of $^{n}J_{HH}$ is aligned with those previously reported in the literature, confirming that the analysis is valid and that a linear conformational bias is likely, based on H-H scalar couplings.

The accuracy and validity of the analysis can be increased by incorporating more NMR parameters into the analysis. As discussed previously, $^{3}J_{CH}$ values also demonstrate a relationship to dihedral angle and can thus provide useful conformational information. On examination of structure **119**, it is clear that two HCCC dihedral angles can be used to provide information regarding the conformation of the backbone. The HCCC dihedral angle between a backbone CH proton and the relevant backbone carbons (i.e. $H_1C_1C_2C_3$) should always be ~60° resulting in a small $^{3}J_{CH}$ value. Additionally, the dihedral angles between the backbone CH protons and the carbon of the CH_3 or CH_2 groups can be examined. For the terminal CH protons (H_1 and H_7) one dihedral angle can be examined, both of which should be ~60°. For H_2 – H_6 two dihedral angles can be examined, one of which should be ~180 ° and one ~60° (**Figure 76**). Thus, provided they can be experimentally determined, a large number of potential $^{3}J_{CH}$ values can be measured to provide more information regarding the conformation of **119** in solution.

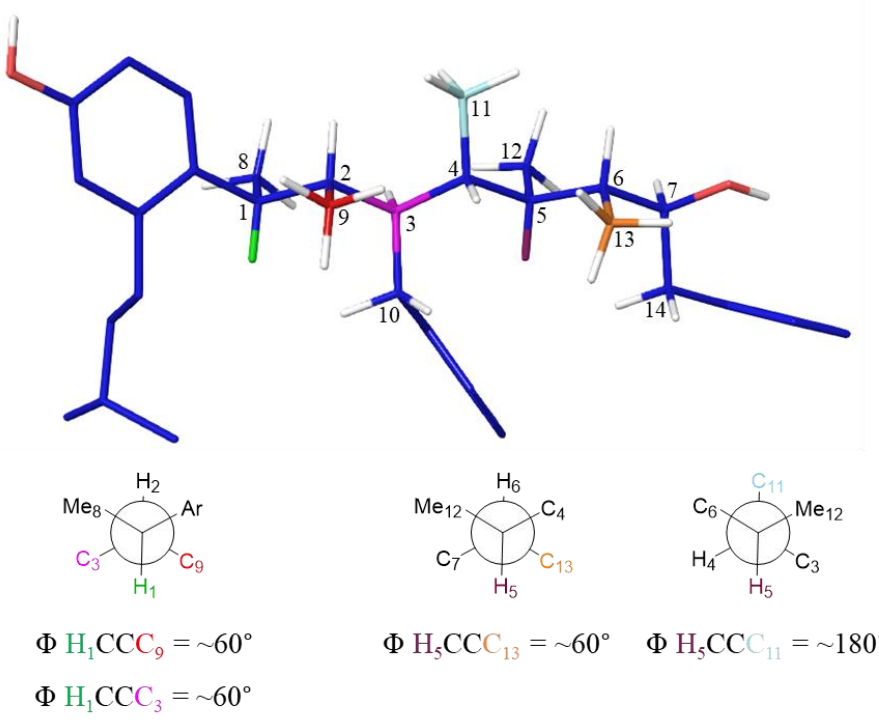


Figure 76. Examples of HCCC dihedral angles that will provide information regarding the conformation of **119**.

The acquisition of $^{3}J_{CH}$ values experimentally by NMR spectroscopy is less straightforward than that for $^{3}J_{HH}$. IPAP-HSQMBC (Inphase-Antiphased heteronuclear single quantum multiple bond correlation) was chosen for the analysis as it has been reported by Butts *et. al.* that the IPAP-HSQMBC is the best option if a global analysis of $^{n}J_{CH}$ values is desired.[99] The IPAP-HSQMBC is a relatively time efficient method as it does not require the acquisition of multiple datasets. It is however challenging to measure couplings of <2 Hz due to the insensitivity of the experiment. The IPAP-HSQMBC proved to be very successful for **119**, with a total of forty-seven values of $^{n}J_{CH}$ measured. Of these, eleven corresponded to $^{3}J_{CH}$ values along the backbone that are of interest for the conformational analysis of **119** (**Table 8**). Overlap in the F_2 dimension prevented the measurement of further $^{3}J_{CH}$ values of interest.

The dihedral angle between H_7 and C_{13} is expected to be >5Hz for a linear conformation. However, a dihedral angle of 4.4 Hz was measured experimentally, with a comparable value obtained from DFT calculations. The dihedral angle in question is affected by the terminal hydroxy group. The effect of electronegative substituents on the value of $^{3}J_{CH}$ has been investigated previously by Bifulco *et. al.* who explored the effect of electronegative substituents on butane to the $^{3}J_{CH}$ values.[277] They found that the presence of electronegative atoms reduced the value of $^{3}J_{CH}$ from the value predicted by Karplus equations. For example, the value of the coupling constant between H_{33}-Me_{32} in the marine natural product Sphinxolide, **180**, was measured to be 4.4 Hz and calculated to be 4.8 Hz, despite a dihedral angle of ~180° between these two groups, caused by the adjacent methoxy group (**Figure 77**).[277]

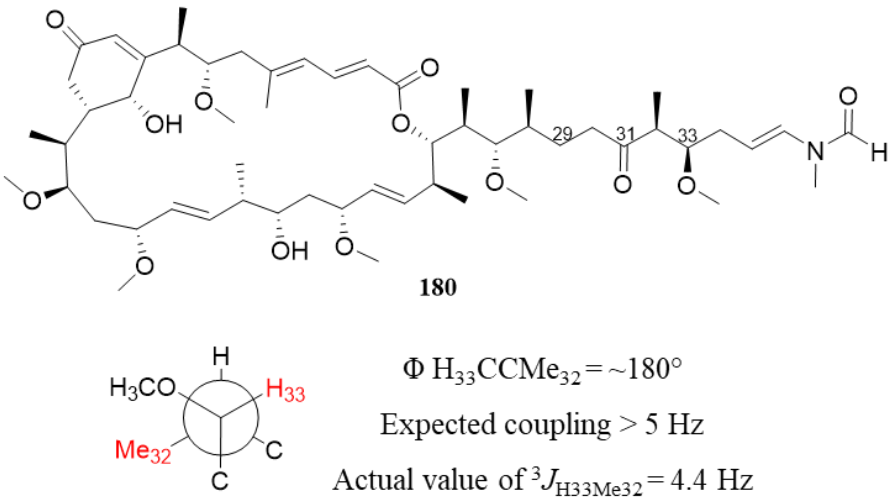
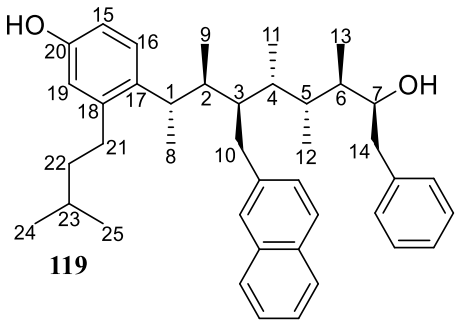


Figure 77. The effect of electronegative atoms on the $^{3}J_{CH}$ coupling constant is clear on examination of the H_{33}-Me_{32} coupling.[277]

The experimentally measured values of $^{3}J_{CH}$ were compared to those calculated using DFT, and an excellent correlation was observed, with a MAD of 0.32 Hz and a SD of 0.29 Hz obtained. This will well within the bounds of MADs and SDs reported in the literature for molecules of similar flexibility. In addition to the eleven values $^{3}J_{CH}$, the remaining thirty-six experimental values that were able to be

measured were also compared to those calculated using DFT. The comparison of all fourty-seven experimental and calculated $^{n}J_{CH}$ values gave a MAD of 0.35 Hz and a SD of 0.48 Hz.

Table 8. The $^{3}J_{CH}$ values along the backbone of the molecule capable of being measured using IPAP-HSQMBC



119

H	C	Relationship	Dihedral Angle (°)	Anticipated Coupling (Hz)	Experimental Coupling (Hz)	Calculated Coupling (Hz)*	Deviation
1	9	*Gauche*	60	Small <3	2.5	1.7	0.7
2	8	*Gauche*	60	Small <3	3.3	3.3	0.0
3	11	*Gauche*	60	Small <3	3.2	2.4	0.8
3	9	*Anti*	180	Large >5	6.1	6.2	0.1
4	12	*Anti*	180	Large >5	5.6	5.5	0.1
4	2	*Gauche*	60	Small <3	2.7	2.2	0.5
5	7	*Gauche*	60	Small <3	2.2	2.4	0.2
6	12	*Gauche*	60	Small <3	2.6	2.6	0.1
6	14	*Anti*	180	Large >5	5.0	5.2	0.2
7	5	*Gauche*	60	Small <3	2.0	1.4	0.7
7	13	*Anti*	180	Large >5	4.4	4.5	0.1
						Mean Absolute Deviation (Hz)	**0.32**
						Standard Deviation (Hz)	**0.41**

*mPW1PW91/6-311g(d,p)

The impressive correlation between experimentally measured values of $^{n}J_{XH}$ (X = H, C) to the Boltzmann averaged values, calculated by DFT, confirms that the DFT calculated Boltzmann populations of the conformers are accurate and agree with the conformational behaviour of **119** in solution, as observed by NMR spectroscopy. **Figure 78** shows a comparison of all sixty-four experimental coupling constants to those calculated by DFT and a clear positive correlation can be seen from the data. The results from the analysis of experimental and calculated scalar couplings, suggest that a linear conformation is preferred in solution. However, scalar couplings are not highly sensitive

to conformer populations. Therefore a more discriminating technique will also be used to explore the conformational behaviour of **119** in solution. The next section will discuss the conformational analysis of **119** using 1D-NOESY.

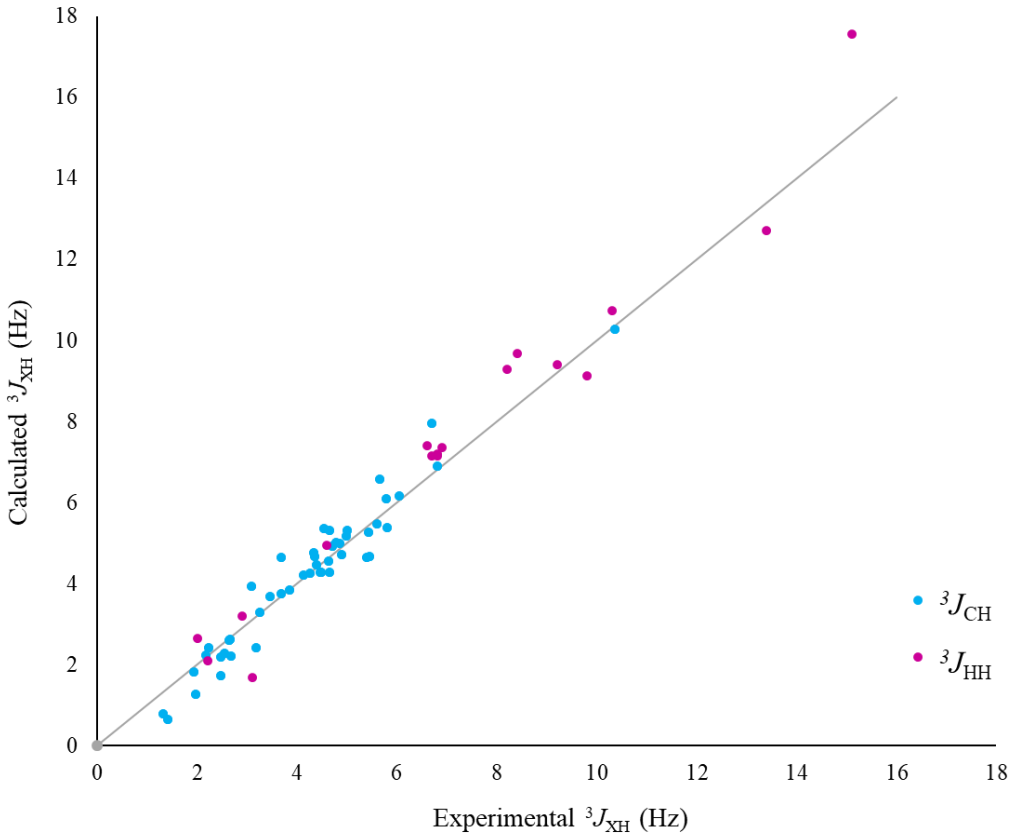


Figure 78. A graph showing the correlation between the experimental and calculated $^nJ_{XH}$ values.

5.4 A Comparison of the Experimental Interproton Distances Obtained from 1D-NOE Spectroscopy to the Boltzmann Averaged Calculated Values

Although scalar coupling constants provide valuable information regarding the conformation of a molecule in solution, they are not sensitive to high energy conformers i.e. those with low Boltzmann populations. For example, if a conformer with a population of <2% was not considered in the conformational analysis, any one Boltzmann-averaged scalar coupling constants would only change by <0.2Hz at most and so it is not likely to highlight this missing conformer. Thus, a technique that can detect weakly populated conformers is desirable in conformational analysis. Fortunately, such a technique exists. Nuclear Overhauser Effect Spectroscopy, or NOESY, provides information regarding the spatial arrangement of protons and can be used to measure quantitative interproton distances for protons <5 Å apart. NOESY is substantially more discriminating against minute conformational contributions than scalar couplings and can be used to identify weakly populated conformers.[278]

The quantitative measurement of interproton distances by NOESY, can be performed using both 1D or 2D NOESY methods. Although 2D- methods provide information on all NOE correlations within the molecule, selective 1D-NOE experiments were chosen for the extraction of quantitative interproton distances, as they are easier to process (phasing and baseline correction) and do not suffer from t_1-noise.

One limitation to a simple selective 1D-NOE experiments is that ~50 Hz separation between ^{1}H resonances in the 1D proton spectrum is required for clean irradiation of the desired peak. Unfortunately, for **119** only three ^{1}H are isolated enough to selectively irradiate. However, the CSSF NOESY addresses this and allows highly selective irradiation of overlapped ^{1}H resonances using chemical shift selective filtering (**Figure 79**). Provided two overlapped ^{1}H peaks have different ^{1}H chemical shifts, and are not strongly coupled, one can be selectively irradiated over the other, regardless of how much their multiplets overlap.[279,280] A pureshift 1D-^{1}H NMR spectrum must be first acquired since the exact chemical shift of the peak of interest must be known. Using CSSF-NOESY, an additional four ^{1}H resonances were able to be selectively irradiated. The seven ^{1}H resonances irradiated yielded a total of twenty-nine NOE correlations that could be accurately analysed.

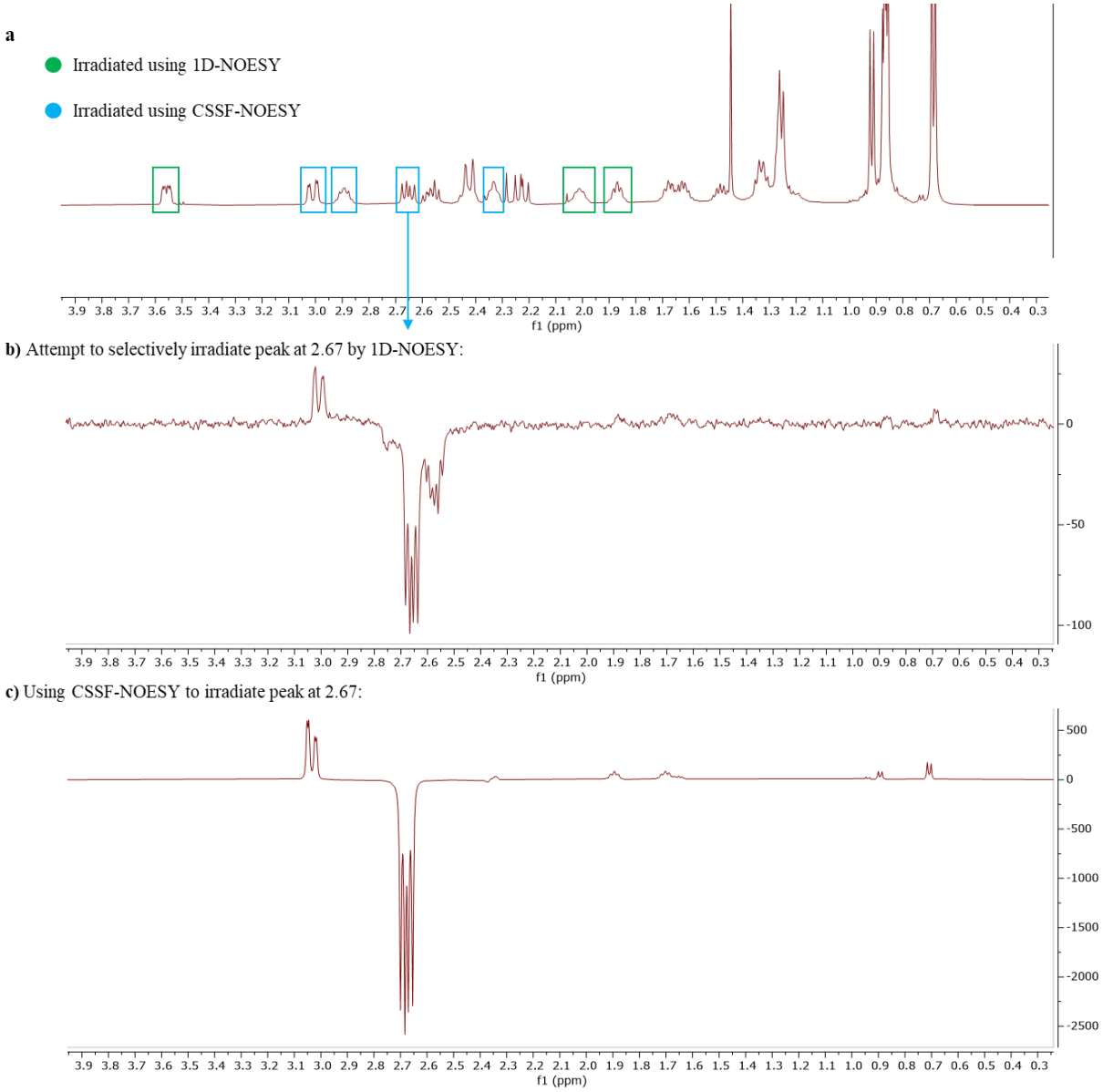


Figure 79. (**a**) The seven ^{1}H resonances irradiated by both 1D-NOE and CSSF-NOESY. The advantage of using CSSF-NOESY in overlapped regions of the spectra is exemplified by the unsuccessful irradiation of the peak at 2.67 ppm by 1D-NOE (**b**) and the successful irradiation of the same peak using CSSF-NOESY (**c**).

Section 1.3.1 of Chapter 1 discussed how NOE signal intensities (η_{NOE}) can be used to derive interproton distances using the PANIC method.[106] The PANIC technique determines accurate interproton distances by comparing the relative NOE intensities from a selective inversion experiment to a reference peak, typically the irradiated peak, that has been set to an arbitrary intensity (usually 1000).[107] Provided that the molecule is in the fast tumbling regime, the ratio of the NOE intensities of two separate nuclei pairs to the ratio of the corresponding interproton distances (raised to the negative sixth power). Thus, using a well-defined reference distance, typically one not affected by conformational dynamics, other interproton NOE-distances can be derived. The preferred choice of reference distance is that between geminal protons of a methylene pair since the distance between the protons (~1.76 Å) is not affected by

changes in the geometries of contributing conformers. Fortunately the methylene pair $H_{10'}$ and $H_{10''}$ of **119** can be selectively irradiated and used as the reference distance (r_{ref}). The NOE for the reference distance is prone to the same errors as every other NOE in the molecule. To correct for this, after calculation of an initial set of NOE-distances for the whole molecule, the reference NOE-distance is subsequently adjusted to maximise the quality of fit between experimental and computed distances. Strictly speaking this means the entire dataset is being used as "reference distances" rather than relying on one, potentially erroneous, datapoint. In this instance, the best fit was found with the reference distance set to be 1.74 Å. A summary of the NOE-interproton distances derived from 1D-NOE and CSSF-NOESY is shown in **Table 9.**

The integration of the irradiated ^1H resonance was set to 1000 and the integration of the NOE correlations were measured. NOE correlations with an integral of <5 were eliminated from the analysis due to inaccuracy in the integration of such weak correlations. Additionally, NOE correlations arising from overlapping ^1H resonances were not considered as the contributions of each ^1H resonance to the NOE intensity cannot be easily deduced. For a CH→CH_3 NOE correlation, the resulting integral of the CH_3 peak was divided by three in order to average the three distances between the CH proton and the three equivalent CH_3 protons, as described in the supporting information from Aggarwal *et. al.*[281] The NOE intensities (η_{NOE}) were converted to interproton distances (r_{NOE}) using the $H_{10'\text{-}10''}$ pair as the reference NOE (η_{ref}) and distance (r_{ref}). For example, the interproton distance experimentally derived for the $H_{10''}$-H_4 distance ($r_{H10'',H4}$) is obtained from the ratio of the $\eta_{H10'',H4}$ intensity to the $\eta_{H10',H10''}$ intensity, multiplied by the reference distance (**Figure 80**).

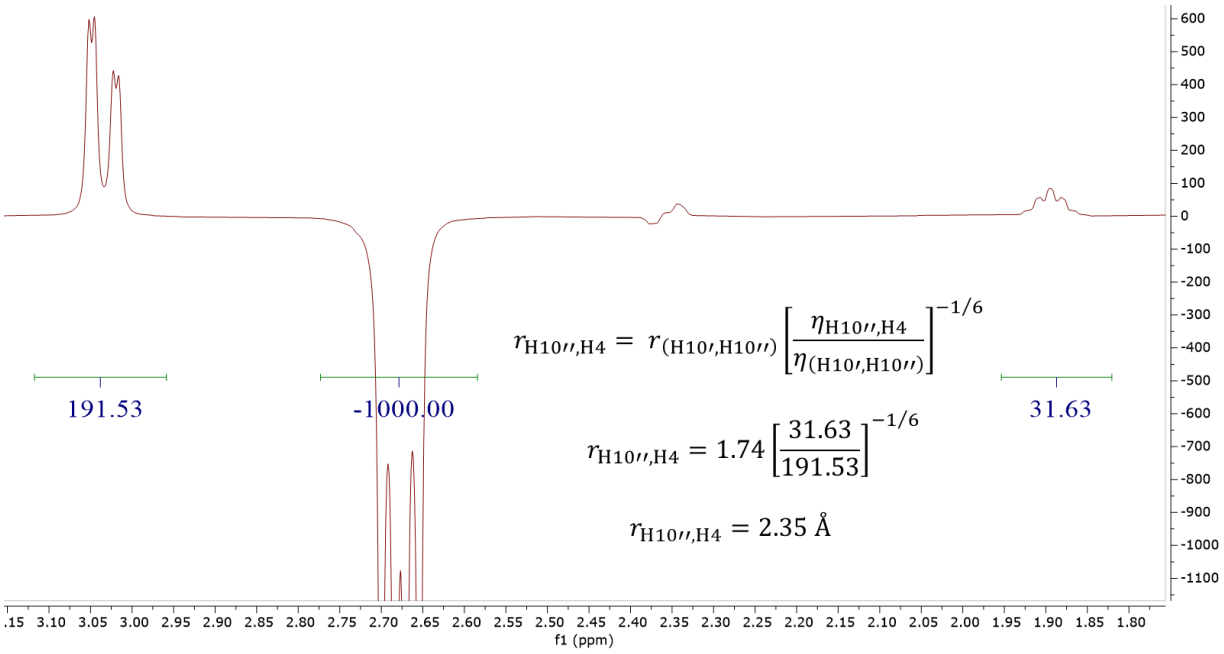
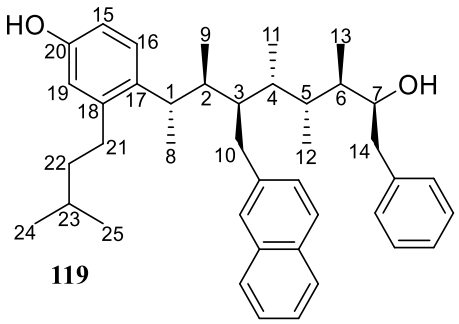


Figure 80. Using 1D-NOESY to experimentally measure the H_4.H_3 interproton distance. The reference distance chosen was that between methylene protons $H_{10'}$-$H_{10''}$.

The experimental NOE-interproton distances can be compared to the Boltzmann averaged interproton distances obtained from DFT. The NOE intensities are very sensitive to conformer populations and a poor correlation to DFT will be observed if the conformer populations are not accurate. For a rigid molecule a MAD of ~3% or 0.1 Å is typically observed if the calculation of conformer populations is accurate.[104] Although molecule **119** displays a conformational bias, it contains a large number of rotatable bonds making the calculation of its conformational landscape much more challenging. As discussed in section 5.2 >17000 conformations of **119** were found within 30 kJ mol^{-1} of the global minima. Approximately 1.5% of these conformers were taken through to DFT calculations resulting in an approximation of the conformational landscape of the molecule rather than an exhaustive analysis of the conformational landscape. This will likely result in a larger deviation between experimental and calculated interproton distances due to the sensitivity of NOE intensities to conformer populations. For flexible molecules, deviations of <5% are considered to be acceptable. Indeed, for the molecules studied by Aggarwal *et. al.*, **10**, **11** a deviation of ~5% for interproton distances were observed.[27]

The experimentally derived interproton distances (**Table 9**, column 4) were compared to the Boltzmann averaged calculated values (**Table 9**, column 5) and a MAD of 3.7% and SD of 5.4% were obtained. This strong agreement between the two datasets further confirms that a linear conformational bias, as predicted by DFT calculations, is preferred.

Table 9. The experimental NOE-interproton distances obtained from 1D-NOE and CSSF-NOESY and the Boltzmann averaged interproton distances, calculated by DFT. The MAD and SD are also shown



119

Irradiated ^{1}H (Ha)	Observed ^{1}H (Hb)	η_{NOE}	Experimental r_{AB} (Å)	Calculated r_{AB} (Å)*	Absolute Deviation (%)
7	14’	17.18	2.60	2.49	4.42
	14’’	6.77	3.04	3.04	0.02
	4	14.04	2.69	3.11	15.66
	13	2.37	3.62	3.59	0.63
	12	12.67	2.74	2.77	1.40
10’	1	60.52	2.11	2.08	1.48
	10’’	182.70	1.75	1.74	0.64
	3	19.99	2.54	2.67	5.28
1	10’	34.86	2.31	2.08	10.13
	20’	40.61	2.25	2.19	2.84
	20’’	21.76	2.50	2.45	2.20
	3	11.73	2.77	2.91	5.03
	2	5.66	3.13	3.02	3.56
10’’	**10’ (ref)**	**191.53**	**1.74**	**1.74**	**0.15**
	4	31.63	2.35	2.41	2.53
3	10’	15.69	2.64	2.67	1.11
	1	13.11	2.72	2.91	7.00
	2	25.06	2.44	2.48	1.54
	4	13.14	2.72	2.83	4.08
	8	18.88	2.56	2.51	1.97
	11	4.27	3.28	3.27	0.19
2	16	48.81	2.19	2.27	4.10
	3	22.77	2.48	2.48	0.07
	8	6.96	3.02	2.93	3.18
	11	15.60	2.64	2.68	1.46
4	7	14.60	2.67	3.11	16.42
	10’’	20.57	2.52	2.41	4.56
	3	11.13	2.80	2.83	1.24
	13	15.17	2.66	2.75	3.49
				Mean Absolute Deviation (%)	**3.67**
				Standard Deviation (%)	**5.42**

*mPW1PW91/6-311g(d,p)

The %deviation between the experimental NOE-interproton distances and the Boltzmann averaged DFT calculated interproton distances is <5%. However, the H_7-H_4 distance, gives an average deviation of 16.0% or 0.41 Å. The experimental distance measured for the H_7-H_4 distance is 2.68 Å, averaged across the two distances measured, whilst the calculated distance is longer at 3.11 Å. This large deviation can be clearly seen when looking at **Figure 81**, which shown the %deviation obtained for all interproton distances.

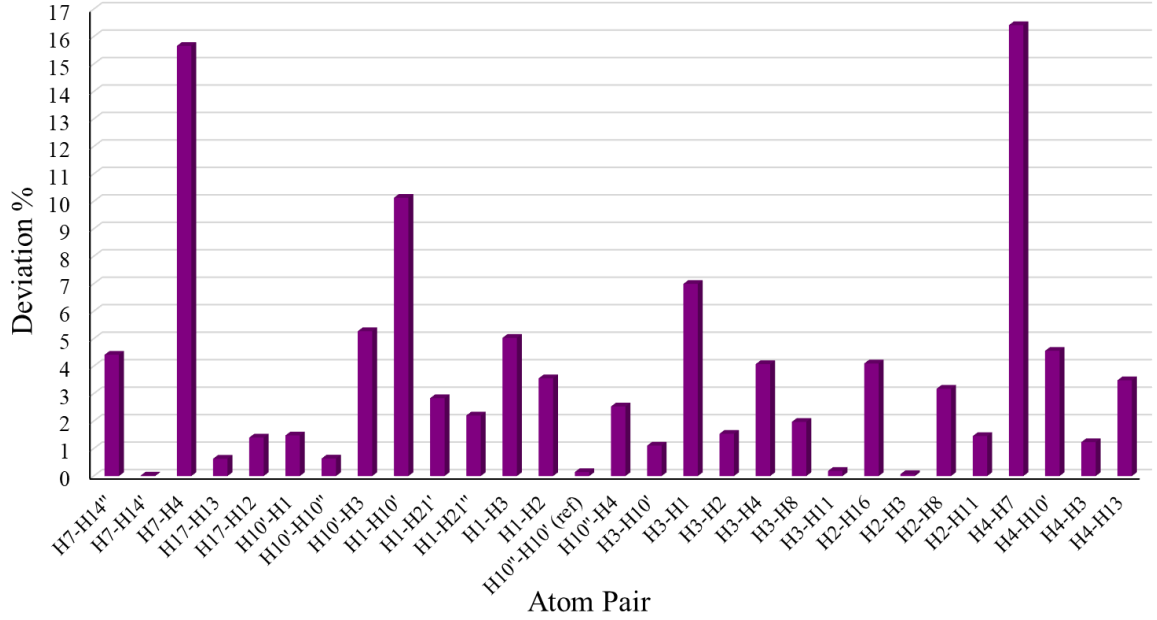


Figure 81. A graph showing the %deviation between calculated and experimental interproton distance obtained for each atom pair investigated by 1D-NOE.

As H_4 and H_7 give well isolated peaks then experimental issues in the accurate integration or assignment of the NOE peak is unlikely. Additionally, very similar experimentally measured values of the H_7-H_4 distance are obtained when either H_7 is irradiated or H_4. This suggests that the population of conformers with a short H_7-H_4 distance are underestimated by DFT. As discussed previously the NOE intensity between two nuclear spins is strongly related to interproton distance, scaled with r^{-6}. This implies that conformers with a shorter value of $r_{4\text{-}7}$ will have a disproportionally large effect of the NOE intensity, even if those conformers are present in low quantities.[102] The distance between H_7 and H_4 for conformers adopting a linear conformer is ~4.75 Å however, the distance between H_7 and H_4 for the next most populated cluster of conformers (see the bubble plot in **Figure 72**) is ~2.20 Å (**Figure 82**). The results from the NOE-interproton distance analysis suggests that the combined population of the second most populated cluster of conformers is >10%. This would require a reduction in the combined population of the linear conformers. However, the change in population is likely to be very small due to the disproportionally large effect that conformers with a short distance have on the NOE intensity.

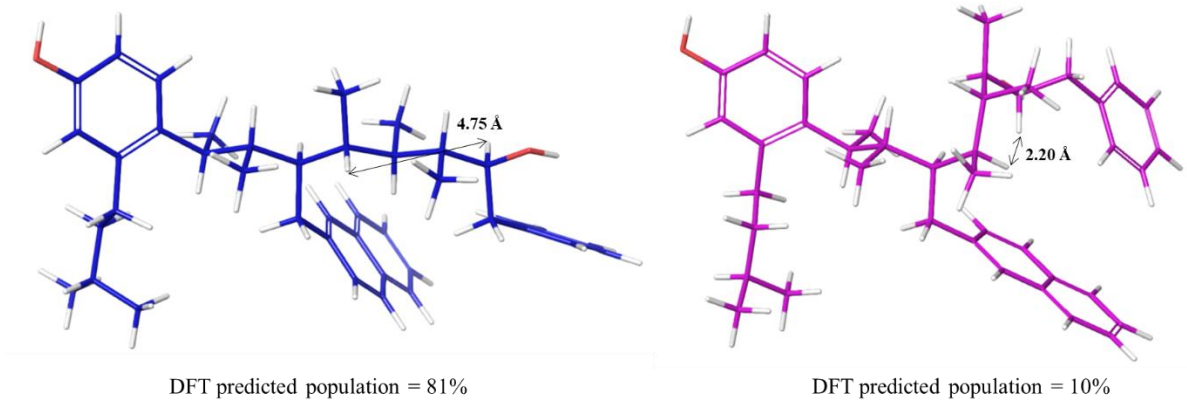


Figure 82. A representative linear conformer with the H_4-H_7 distance shown and a representative conformer from the second most populated cluster of conformers with the shorter H_4-H_7 distance shown.

The other ^{1}H resonance that presents a much large deviation between experiment and calculation is the H_1-$H_{10''}$ interproton distance, with a deviation of 10.1%. The corresponding $H_{10''}$-H_1 distance however presents a typical deviation of 1.5%. This suggests that experimental reasons are behind the large deviation observed for the H_1-$H_{10''}$ interproton distance. Examination of the CSSF-NOEY spectra of H_1 reveals that $H_{10''}$ is separated from H_1 by ~50 Hz. It is possible that the selective irradiation of H_1 was not achieved.

In summary, the interproton distances obtained from 1D-NOE and CSSF-NOESY experiments correlated well with those calculated by DFT. This suggests that the linear conformational bias observed from DFT calculations is also observed experimentally in solution.

5.5 A Comparison of the Experimental ^{1}H and ^{13}C Chemical Shifts to the Boltzmann Averaged Calculated Values

Alongside scalar couplings and quantitative interproton distances, ^{1}H and ^{13}C chemical shifts can be used to obtain conformational information about a structure. Chemical shifts are very sensitive to their surrounding chemical environment, and the conformational bias of a molecule can sometimes be deduced from the 1D-^{1}H or ^{13}C spectrum. For example, as discussed in section 1.1.3 of Chapter 1 a team of scientists at AstraZeneca used ^{1}H chemical shifts to confirm that the free ligand conformation of potential drug candidates reflected the bound conformation of the molecules. Of all of the NMR parameters, chemical shifts are the easiest to measure and require the least experimental time.

The experimental chemical shifts, both ^{1}H and ^{13}C were measured directly from the corresponding 1D-spectrum. The acquisition of a pure-shift ^{1}H spectrum, which removes scalar coupling to other ^{1}H resonances, was critical in obtaining accurate values of δ_H due to peak broadening and overlap. Following the experimental acquisition of both δ_H and δ_C, the DFT calculated values was performed. The DFT NMR GIAO calculations calculate magnetic shielding tensors (σ) which are converted to a chemical shift by referencing the magnetic shielding tensors to those calculated for trimethyl silane (TMS) at the same level of theory in the same solvent, using **Equation 8**.

$$\delta_{calculated} = \frac{\sigma^{TMS} - \sigma^{calculated}}{1 - \sigma^{TMS}\ x\ 10^{-6}}$$

Equation 8. The conversion of magnetic shielding tensors to chemical shifts by referencing to TMS.

The calculated chemical shifts were then scaled according to **Equation 9**:

$$\delta_{scaled} = (\delta_{calculated} - intercept)/slope$$

Equation 9. The scaling of calculated chemical shifts.

Where the intercept and the slope were obtained by plotting the calculated chemical shifts ($\delta_{calculated}$) against the experimental chemical shifts. The linear scaling has been shown to reduce systematic errors and is performed routinely in the comparison of experimental chemical shifts against calculated chemical shifts.[110,135] After scaling, the corrected calculated chemical shifts were compared to the experimental chemical shifts and a MAD of 0.11 ppm and SD of 0.16 ppm was observed for δ_H and a MAD of 1.14 ppm and SD of 1.46 ppm was observed for δ_C. The full tables of the experimental and computed chemical shifts are given in section 9.6.4 of Chapter 9. Again, the MAD and SD align with previously obtained MAD and SD deviations in the literature for flexible molecules, further confirming that the conformation of **119** in solution behaves as predicted computationally, suggesting a strong bias towards a linear conformation.

5.6 A summary of the Conformational Analysis of 119

The conformation of **119** in solution has been explored extensively through a hybrid computational and NMR spectroscopy approach. This has involved an experimental examination of scalar coupling constants, chemical shifts and quantitative interproton distances using 1D-NOESY followed by a comparison of the experimental data to that calculated using DFT calculations. An excellent correlation between computational and experimental data was observed for all NMR parameters (**Figure 83**), and the MADs and SDs obtained are of comparable quality to literature reports for molecules of similar size and flexibility (**Table 10**). This confirms the computational prediction that the conformational bias of molecule **119** in solution is towards a linear conformation.

Table 10. A summary of the comparison of the experimental NMR parameters to the calculated NMR parameters for **119**

NMR Parameter	MAD	SD
$^{n}J_{HH}$	0.71 Hz	0.85 Hz
$^{n}J_{CH}$	0.35 Hz	0.48 Hz
Interproton Distances	3.7%	5.4%
δ^{1}H	0.11 ppm	0.16 ppm
δ^{13}C	1.14 ppm	1.46 ppm

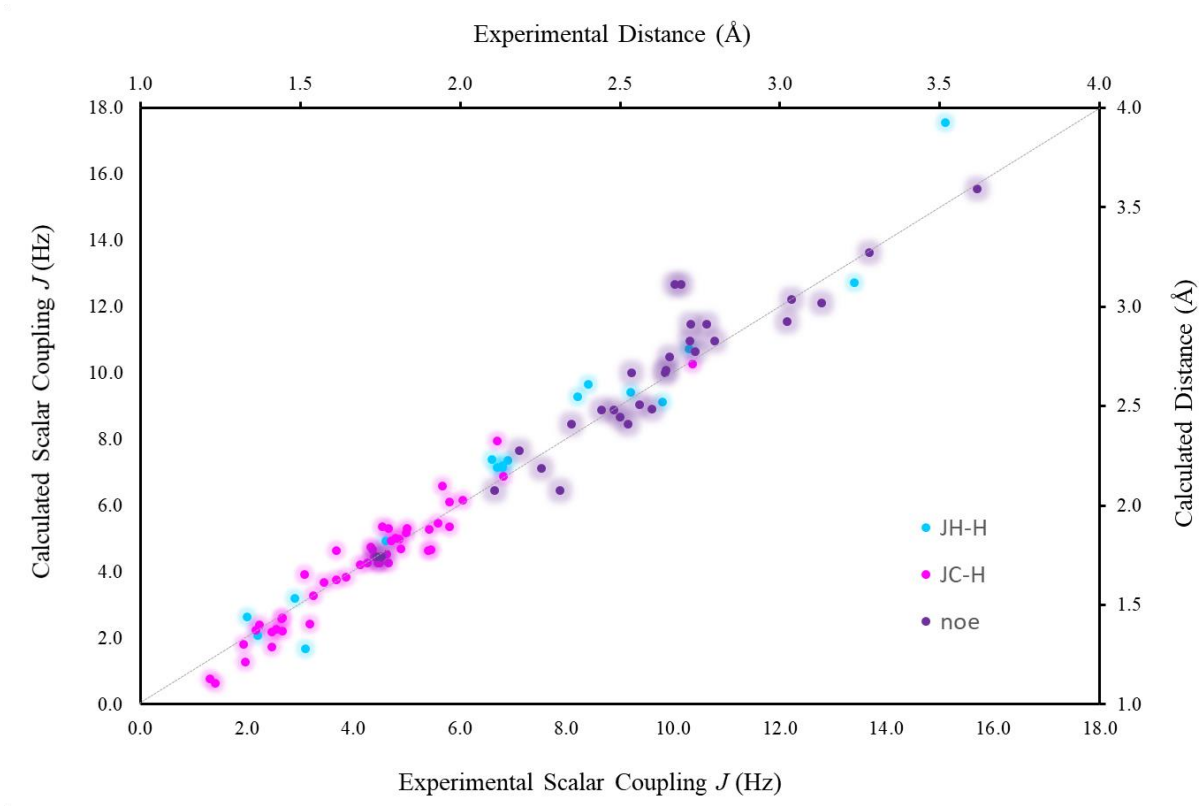


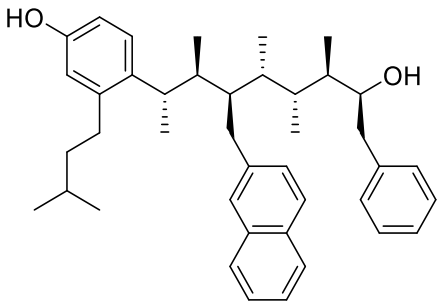
Figure 83. A graph to show the correlation between the experimentally derived scalar coupling constants and interproton distances to those calculated by DFT.

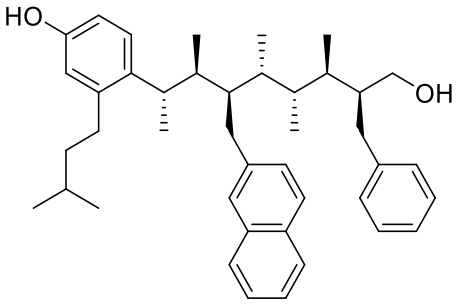
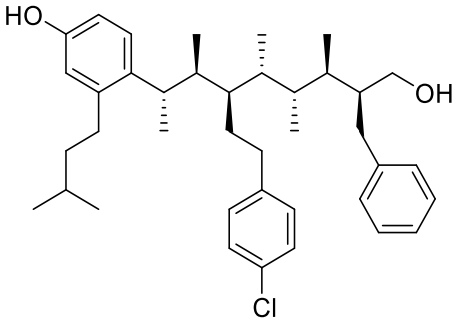
5.7 An Examination of Structures 155, 157 and 163.

This section will investigate the conformation of **155**, **157** and **163** computationally by MM conformational searching and experimentally through a qualitative examination of their $^{3}J_{HH}$ coupling constants.

A full conformational analysis of all structures is not necessary as MM conformational searching suggests that structures **155** and **163** also demonstrate a linear conformational bias. The additional methylene group (**155**) and the substitution of the 'hot-spot' group from naphthyl (**119**) to the *para*-chloro substituted phenyl group (**163**) has a minimal effect on the conformational bias and the conformation is still controlled by *syn*-pentane interactions. For molecule **119**, a very small change in the total percentage of conformers adopting a linear conformation was observed when moving from MM to DFT, although the populations of the individual conformers changed considerably. Therefore, the conformational bias of **155** and **163** as predicted by MM can be assumed to be reliable. $^{3}J_{HH}$ coupling constants will however, confirm that the molecule is displaying a linear conformational bias. A conformational search of the PEG derivative, **157**, cannot be performed as simulation of the PEG chain in implicit solvent is not accurate and will require considerable computation time due to the number of rotatable bonds. However, the values of $^{3}J_{HH}$ coupling constants measured from the NMR spectra of **157** should indicate what the conformational bias is in solution. A summary of the conformational search results for **119**, **155** and **163** are shown in **Table 11**.

Table 11. A summary of the conformational search results for compounds **119**, **155** and **163**. The bubble plots are provided in section 9.1.3 of chapter 9.

Structure	Number of Conformers within 21 kJ mol^{-1}	Number of Conformers after Redundant Conformer Elimination	% of Linear Structures.
 **119**	4137	270	77%

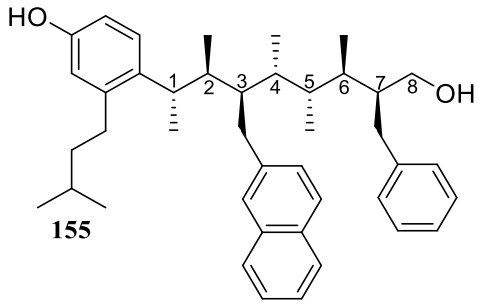
 **155**	9577	200	72%
 **163**	6277	188	75%

As discussed previously, the values of $^{3}J_{HH}$ provide an indication as to the conformational bias of a molecule due to their relationship with dihedral angle, described by the Karplus equation. For molecules, **155**, **157**, and **163** SpinSimulation had to be used to extract values of $^{3}J_{HH}$ due to broadening and overlap of the peaks precluding the measurement of these values directly from multiplets in the 1D ^{1}H NMR. This proved very successful for **119** and all values of $^{3}J_{HH}$ were obtained for the backbone HCCH dihedrals.

For structure **155**, many of the ^{1}H resonances are significantly overlapped and broadened to a greater extent than structure **119**. This makes simulation of these peaks very challenging. It is clear from the spectrum in **Figure 84** that the ^{1}H resonances for H_2, H_3, H_5, H_6, H_7 are all either too overlapped or broad to obtain reliable values of $^{3}J_{HH}$. However, the ^{1}H resonances for H_1 and H_4 were able to be simulated and the values of $^{3}J_{HH}$ obtained are consistent with a linear conformational (

Table 12). These coupling constants, along with the MM conformational search, suggest that the additional methylene group has not affected the overall conformational bias of the molecule towards a linear conformation.

Table 12. The three values of $^{3}J_{HH}$ obtained for **155**. The experimental coupling constants matched the anticipated coupling constants for a linear conformational bias.



155

Proton	Proton	Relationship	Dihedral Angle (°)	Anticipated Coupling (Hz)	Experimental Coupling (Hz)
1	2	*Anti*	180	Large >9	9.9
3	4	*Anti*	180	Large >9	2.0
4	5	*Gauche*	60	Small <5	8.7

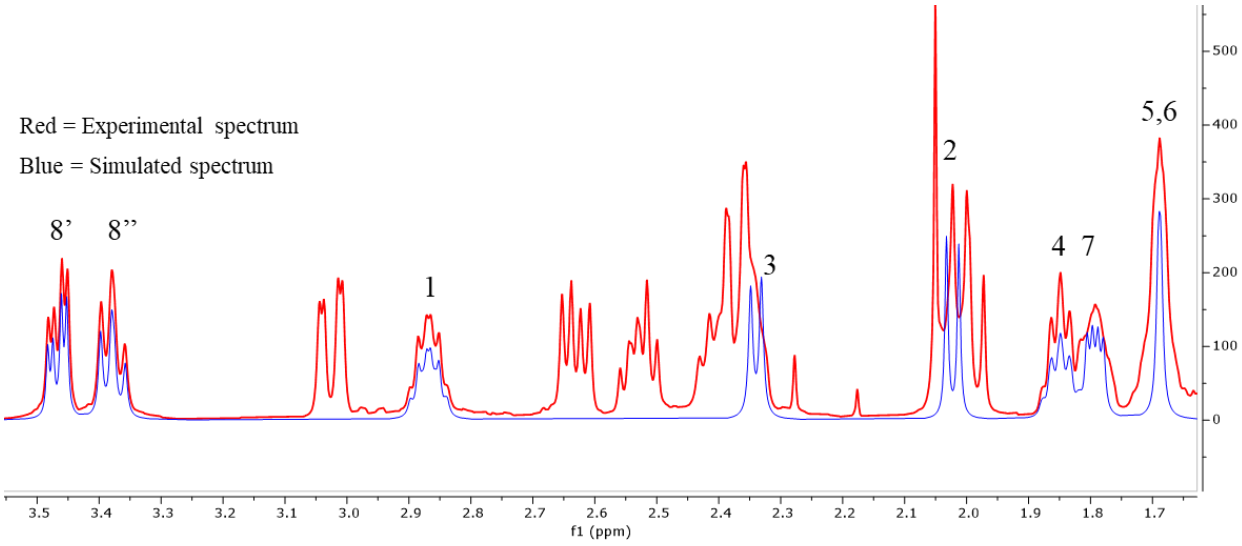
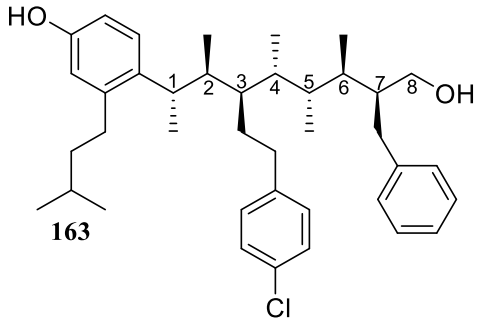


Figure 84. The experimental spectrum of **155** (red) and the simulated spectrum (blue). Only three ^{1}H resonances were able to be reliably simulated, however this did yield three values of $^{3}J_{HH}$ for the backbone HCCH dihedral angles.

For molecule **163**, the peaks in the 1D-^{1}H spectrum were also significantly broadened and overlapped. Again, this led to challenges in the SpinSimulation and only three of the CH backbone peaks were able to be accurately simulated (**Figure 85**). The peaks corresponding to H_1, H_2 and H_6 were successfully simulated and the four coupling constants obtained match with the anticipated coupling constants for a linear conformational bias (**Table 13**). The four experimental coupling constants along with the

computational MM conformational search suggest that changing the hot-spot group from a naphthyl to a *para*-chloro phenyl group does not affect the conformational bias and a linear conformational is preferred for **163.**

Table 13. The four values of $^{3}J_{HH}$ obtained for **163**. The experimental coupling constants matched the anticipated coupling constants for a linear conformational bias.



163

Proton	Proton	Relationship	Dihedral Angle (°)	Anticipated Coupling (Hz)	Experimental Coupling (Hz)
1	2	*Anti*	180	Large >9	10.3
2	3	*Gauche*	60	Small <5	1.3
5	6	*Anti*	180	Large >9	9.9
6	7	*Gauche*	60	Small <5	2.6

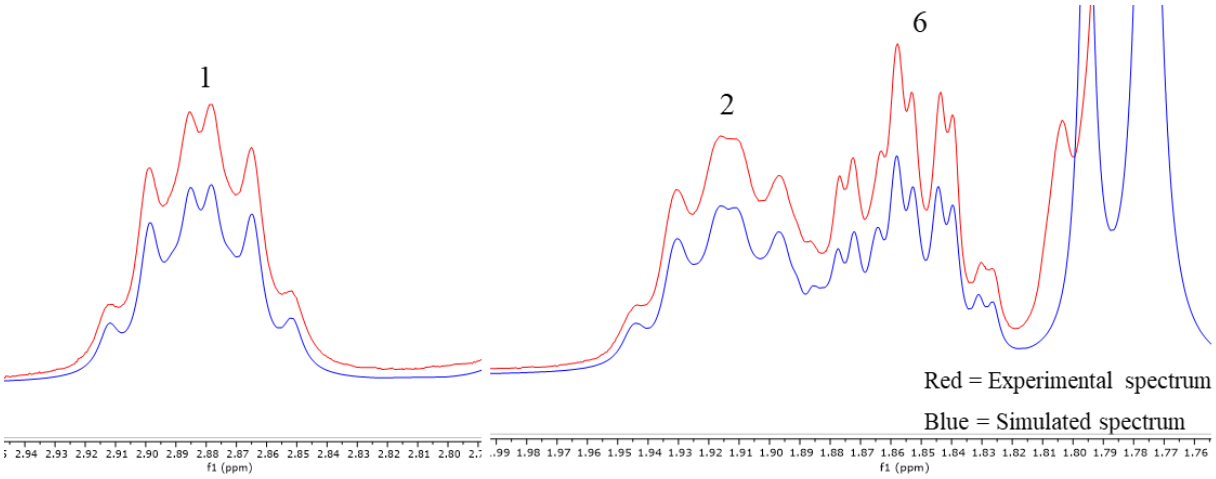
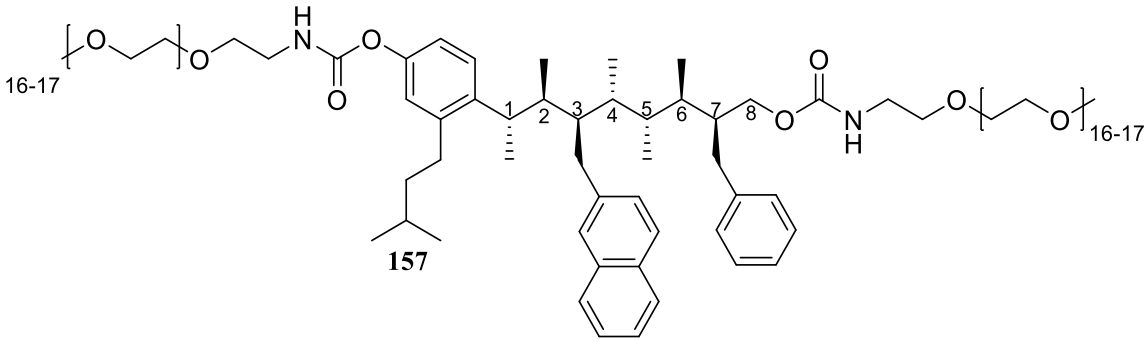


Figure 85. The three ^{1}H resonances of **163** that were successfully simulated using spin simulation.

The PEG derivative **157** was the most challenging to investigate with the highest degree of spectral overlap and peak broadening observed. In addition, only a small amount of material was obtained which, for a molecule with a molecular weight in excess of 2000 Da, resulted in a sample of low concentration. However, all of the backbone CH peaks were able to be assigned. The distribution of the ^{1}H resonances is similar to that observed for its precursor **155**, with the ^{1}H resonances occurring at a similar chemical shift. This is an early indication that no major conformational change has occurred on addition of the

PEG chains, as chemical shifts are very sensitive to their surrounding chemical environment. However, three backbone CH peaks were able to be simulated with some success to provide an estimate for the couplings along the backbone (**Figure 86**). The four values of $^{3}J_{HH}$ along the backbone that could be obtained are shown in **Table 14**. Again, the experimental values of $^{3}J_{HH}$ match with the expected values for a linear conformational bias. This result suggests that the addition of the PEG chain does not disrupt the conformation of the **157**, and a linear conformational bias is still observed.

Table 14. The four values of $^{3}J_{HH}$ obtained for **157**. The experimental coupling constants matched the anticipated coupling constants for a linear conformational bias.



Proton	Proton	Relationship	Dihedral Angle (°)	Anticipated Coupling (Hz)	Experimental Coupling (Hz)
1	2	*Anti*	180	Large >9	10.1
4	5	*Gauche*	60	Small <5	2.4
5	6	*Anti*	180	Large >9	9.7
6	7	*Gauche*	60	Small <5	2.8

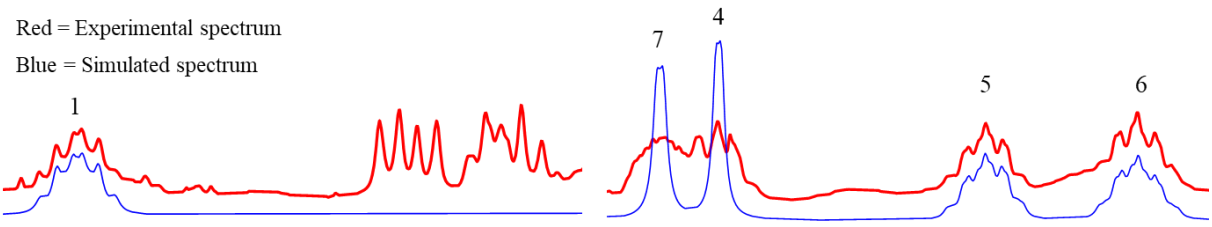


Figure 86. The experimental spectrum of **157** (red) and the simulated spectrum (blue). Only three ^{1}H resonances were able to be reliably simulated.

In summary, all of the molecules investigated within this study show a linear conformational bias regardless of the groups projecting from the conformationally controlled backbone or the size and nature of the end groups. This demonstrates the power of the *syn*-pentane interaction at controlling molecular conformation.

Chapter 6: Analysing the Binding of the Designed p53 Mimetic to Mdm2

The binding of the designed ligands to Mdm2 needs to be examined experimentally to confirm that they act as α-helix mimetics and bind to a pocket on a protein that is typically occupied by an α-helix of a second protein. There are many biophysical methods to study protein–ligand interactions including fluorescence spectroscopy, mass spectrometry, isothermal titration calorimetry and ultraviolet/visible spectroscopy. However, one of the most powerful methods to study protein–ligand binding is NMR spectroscopy, as it provides information on the thermodynamic, kinetic and structural features of protein–ligand binding.[282] There are several NMR methods to study protein–ligand interactions, some of which observe changes to the protein whilst others observe changes to the ligand. For example, Saturation Transfer Difference (STD) detects areas of the ligand interacting with the protein through magnetisation transfer whilst ^{1}H-^{15}N HSQC spectroscopy detects perturbation of chemical shifts of the protein upon binding of a ligand. Protein observed methods are often preferred as it is possible to directly observe the fully saturated protein. An advantage of ligand observed methods is that isotopically enriched proteins do not need to be expressed.[283]

6.1 The Binding of the Designed p53 Mimetics to Mdm2 using ^{1}H-^{15}N HSQC spectroscopy

Studying protein–ligand interactions by NMR spectroscopy is a favourable choice as information regarding the strength and location of binding can be obtained. Of all the NMR experiments available, one of the most widely adopted is the ^{1}H-^{15}N HSQC spectroscopy.[284] The method is simple to perform, the data is straightforward to analyse and on a high field spectrometer a 2D-HSQC can be acquired in ~30 minutes at protein concentrations as low as 50 μM. It is also one of the most effective methods of analysing weak protein–ligand or protein–protein interactions due to the high sensitivity of the experiment.[285] The ^{1}H-^{15}N HSQC is typically performed on ^{15}N labelled proteins due to the low natural abundance of the magnetically active ^{15}N nuclei (~0.4%). In a ^{1}H-^{15}N HSQC spectrum a correlation is seen between the amide proton and its attached nitrogen in the peptide bond, thus each amide in the protein provides a correlation in the 2D-spectrum (Figure 87).[286] Typically, especially if the protein is folded, each amino acid will give a unique peak and little to no overlap will be observed. The binding of a ligand to the protein will cause changes in either the chemical shift, peak intensity or linewidth of

the correlations in the 2D-HSQC spectrum and these can be tracked with increasing ligand concentration to plot binding curves, from which the dissociation constant can be derived.[287] The most commonly observed NMR parameter is the chemical shift perturbation (CSP) upon ligand binding.[288]

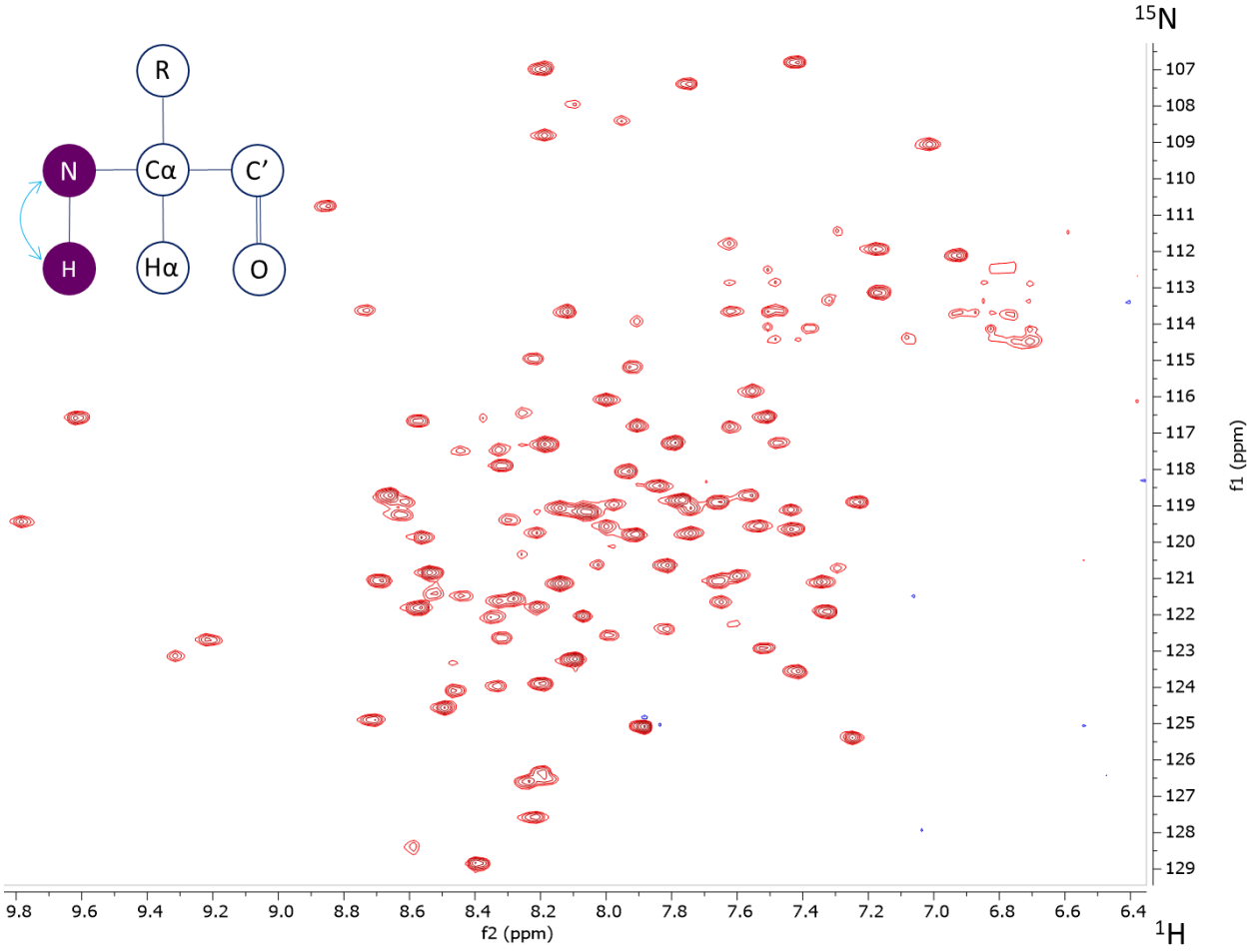


Figure 87. The ^{1}H-^{15}N HSQC spectra of Mdm2 with the polarization scheme for the experiment shown.

CSP's are caused by either changes in the chemical environment of the amino acid caused by interaction of the amino acids with the ligand, or by a conformational change to the protein upon binding of the ligand (**Figure 88**).[289] In the former case, CSP's will only be observed for residues interacting with the ligand, and thus information regarding the location of binding can be derived from the amino acids that have experienced a CSP. In the latter case, an entirely different conformation of the protein could be selected by the ligand, resulting in an entirely different 2D-spectrum, or a small conformational change of a region of the protein can result, resulting in CSP's of just these residues.[290] More often than not, a combination of factors are responsible for the observed CSP's, some may be from interactions with the ligand while others may be caused by a conformational change in this region of the protein.

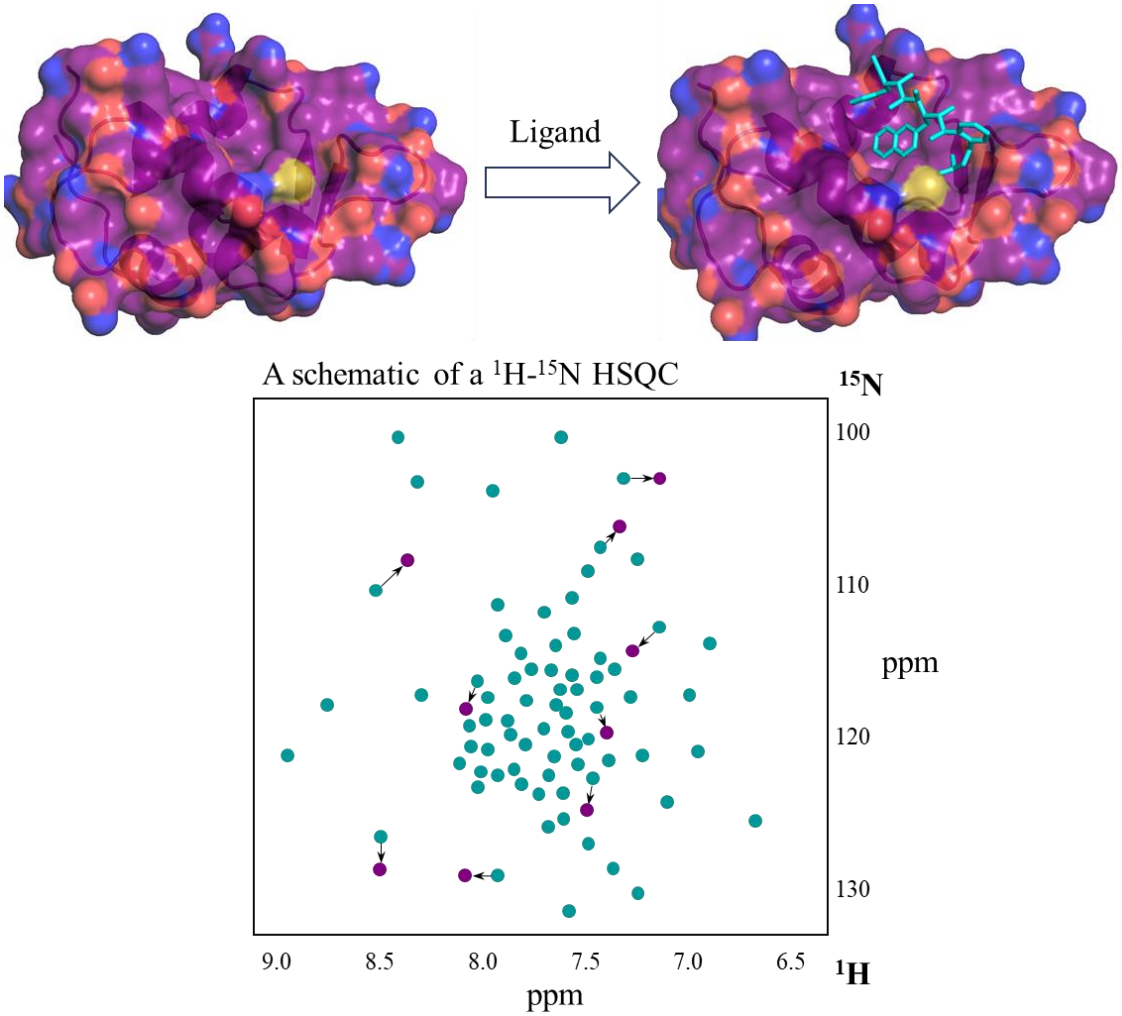


Figure 88. A schematic to demonstrate the CSP effect of addition of a ligand to a ^{1}H-^{15}N HSQC spectrum. On addition of a ligand, CSP's will be observed for amino acids interacting with the ligand if the ligand binds. For example, in the schematic shown above, the green peaks correspond to those of *apo*-Mdm2. Upon binding of the ligand certain residues shift, giving rise to the purple peaks. PDB ID: 1YCR

Using the CSP's it is possible to identify the location of binding, provided the protein has been able to be confidently assigned. However, the ^{1}H-^{15}N HSQC experiment can also provide information regarding the kinetics and thermodynamics of protein–ligand binding. An examination of the CSP's with increasing ligand concentration can help to elucidate the exchange kinetics of protein–ligand complex formation, which provides an indication regarding the strength of binding.[291]

The reversible interaction of a ligand [L] with a protein [P] to form a complex [PL] can be characterised by a second order exchange process governed by the rate constants for association (k_{on}) and dissociation (k_{off}) (**Equation 10**).[292]

$$[P] + [L] \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} [PL]$$

Equation 10. The relationship between free ligand [L] and protein [P] and a protein:ligand complex [PL]

If the exchange rate between free and bound ligand is slow on the chemical shift timescale, then as the ligand is titrated in, the chemical shift corresponding to the free protein gradually decreases while that of the saturated complex gradually increases at the same rate (**Figure 89a**). In the slow exchange regime, k_{off} is slower than the difference in Hz of the chemical shift of the free protein and the saturated complex ($k_{off} \ll \Delta\delta$). Slow exchange kinetics usually indicate a strongly bound ligand. If however, the exchange rate is fast on the chemical shift timescale, then the chemical shift of the free protein gradually shifts from its original position towards the chemical shift of the saturated complex (**Figure 89b**). In the fast exchange regime, k_{off} is faster than the difference in Hz of the chemical shift of the free protein and the saturated complex ($k_{off} \gg \Delta\delta$). Fast exchange kinetics usually indicate a weaker binding ligand. It is also possible for k_{off} to be roughly equal to the chemical shift difference in Hz of the chemical shift of the free protein and the saturated complex ($k_{off} \sim \Delta\delta$). This is known as the intermediate exchange regime and in it, the peaks shift from that of the free to the bound whilst also broadening.[293]-[294] Therefore, on examination of the CSP's following ligand binding it should be possible to identify which exchange regime the system is operating under and indicate the strength of the protein–ligand interaction.

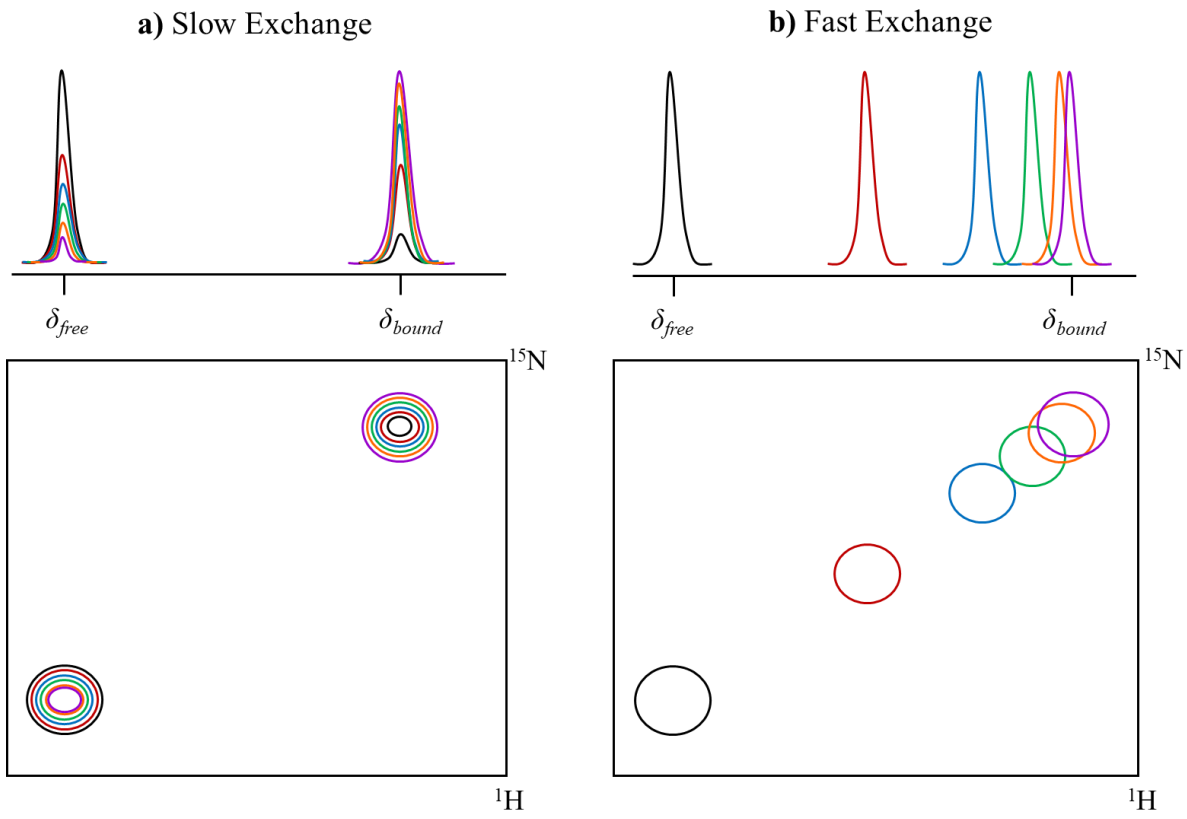


Figure 89. The dependence of the NMR peaks in a ^{1}H-^{15}N HSQC on the exchange rate between protein and ligand. In the fast exchange regime, as ligand concentration is increased, the peaks shift from the free (black) to the bound (purple). In the slow exchange regime, as ligand concentration is increased the free peak (black) decreases in intensity as that of the bound (purple) increases. This figure has been reproduced from reference [290]

Using ^{1}H-^{15}N HSQC spectroscopy, it is also possible to calculate a value for the dissociation constant, K_d, which is equal to the ratio of the rate constant for dissociation (k_{off}) to the rate constant of association (k_{on}), at equilibrium. For systems in the fast exchange regime, the CSP can be monitored as [L] increases to plot a binding curve and calculate the K_d of the protein–ligand interaction (**Figure 90**). Although it is possible, it is much more challenging to measure the K_d of systems in the slow exchange regime as a change in peak intensity is monitored as [L] is increased and this is experimentally much more challenging.[283]

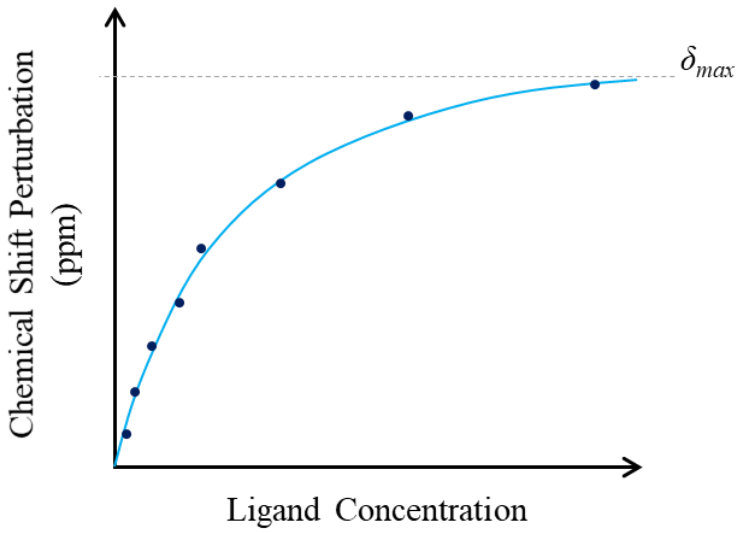


Figure 90. A schematic of a typical binding curve showing the change in CSP as ligand concentration increases. The experimentally measured data points are shown as circles in dark blue and the fitted binding curve is shown as a light blue curve.

The CSP's measured at different ligand concentrations can be used in a nonlinear least-square fitting to calculate K_d using **Equation 11**. This equation assumes the fast regime is operational and that a single binding site is occupied.[295]

$$\Delta\delta_{obs} = \Delta\delta_{max}\frac{\left\{([P]_t + [L]_t + K_d) - [([P]_t + [L]_t + K_d)^2 - 4[P]_t[L]_t]^{\frac{1}{2}}\right\}}{2[P]_t}$$

Equation 11. Using CSP's to estimate K_d.

Where $\Delta\delta_{obs}$ is the change in the observed shift from the shift of the *apo*-protein, $\Delta\delta_{max}$ is the maximum chemical shift change on saturation of the protein with ligand, $[P]_t$ is the total protein concentration and $[L]_t$ is the total ligand concentration. In practice it is very challenging to reach $\Delta\delta_{max}$ due to solubility limitations of the ligand so $\Delta\delta_{max}$ is typically obtained as part of the fitting procedure. The extrapolation of the measured data to estimate $\Delta\delta_{max}$ using nonlinear least-square fitting is one of the major sources of error in the estimation of K_d by CSP's.[291] To calculate the observed CSP, which is contributed to by two different nuclei, an average Euclidean distance is calculated using **Equation**

12. In the Euclidean distance, the chemical shift of the ^{15}N nuclei is scaled by a scaling factor, α. Scaling factors of 0.1[296] to 0.45[297] have been reported and are typically calculated from the ratio of the chemical shift range (in ppm) of ^{1}H to that of ^{15}N.[291] It is also possible to use a scaling factor specific to each amino acid as some amino acids show much greater chemical shift ranges than others.

$$\delta_{obs} = \sqrt{\frac{1}{2}[\delta_H^2 + (\alpha\delta_N^2)]}$$

Equation 12. The calculation of Euclidean distance.

6.2 The ^{1}H-^{15}N TROSY Spectra of Mdm2

The binding of **157** to Mdm2 was explored using ^{1}H-^{15}N HSQC spectroscopy. To achieve this, ^{15}N-labelled Mdm2 was required. ^{15}N-Mdm2 has been expressed numerous times in the literature, with varying lengths of construct used.[265,298–308] It has been noted that Mdm2 is prone to aggregation, and as a result specific pHs, salt concentrations and additives are required to try to stabilise the protein.[298,300] The full details of protein expression and purification are provided in section 9.7.1 of Chapter 9.

A fragment of the N-terminal domain of Mdm2 (residues 27-127) was expressed in the *E.coli* BL21(DE3) strain in M9 minimal media with $^{15}NH_4Cl$ as the nitrogen source. A hexahistidine tag was added to the N-terminus of the Mdm2 construct to aid with purification using immobilised metal affinity chromatography (IMAC). However, all attempts to cleave the His_6-tag with TEV protease were unsuccessful. A report by Vassilev *et. al.* explored the binding of a small molecule inhibitor of Mdm2 by ^{1}H-^{15}N HSQC, using His_6-tagged Mdm2.[304] Therefore perform the analysis of Mdm2 and the binding of the designed ligands was performed using the His_6-tagged Mdm2. Following a second purification by size exclusion chromatography (SEC) using a HiLoad 26/60 Superdex 75 prep grade column, MALDI mass spectroscopy was performed which confirmed that full ^{15}N labelling had occurred. With a purified ^{15}N-labelled sample of Mdm2 in hand, the binding of the designed ligands to Mdm2 by ^{1}H-^{15}N HSQC spectroscopy can be explored.

Although the ^{1}H-^{15}N HSQC is a very powerful experiment for studying protein structure and protein-ligand interactions by NMR, the TROSY derivative of the experiment typically provides much better resolution and sensitivity. It achieves this by selecting the slowest relaxing component of a multiplet in a decoupled HSQC caused by the constructive interference of two relaxation pathways, dipolar-dipolar coupling and chemical shift anisotropy. This reduces the T_2 relaxation rate, leading to sharper lines. The TROSY is typically acquired with a larger number of scans than a corresponding HSQC, since sensitivity is lost when only one component of a multiplet is selected. A ^{1}H-^{15}N TROSY of Mdm2 was acquired at a concentration of 130 μM on a 700 MHz spectrometer fitted with a 1.7 mm inverse triple

resonance micro-cryo probe. The protein was solubilised in a 10 mM $NaHPO_4$ solution with 150 mM NaCl at pH 7.4. The acquired TROSY was compared to reported HSQC or TROSY spectra of Mdm2 published in the literature[249,298,300] which confirmed that the expressed Mdm2 had folded correctly and that the majority of resonances were present (**Figure 91**).

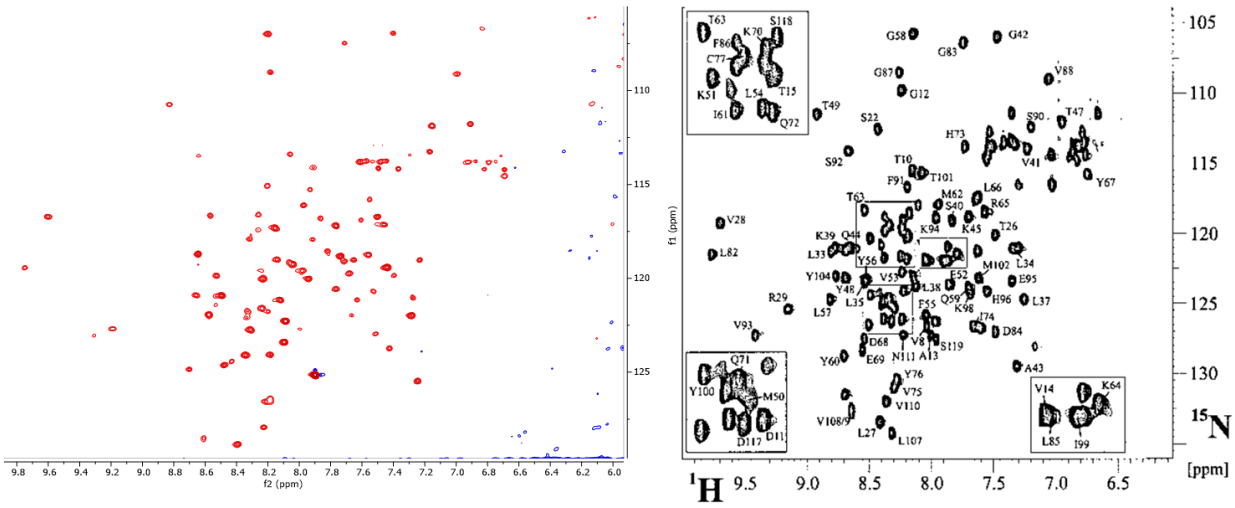


Figure 91. The acquired ^{1}H-^{15}N TROSY of Mdm2 (right) and the ^{1}H-^{15}N HSQC acquired by Holak *et. al.* The graph on the right has been reproduced from reference [300]

Mdm2 is prone to aggregation and the previous conditions used for solvation of Mdm2 resulted in aggregation after 48 hrs. It was found that solubilising the protein in 20mM Tris base with 250 mM NaCl and 1 mM TCEP at pH 7.5 was sufficient at preventing aggregation of Mdm2. A ^{1}H-^{15}N TROSY of Mdm2 in the revised solvation conditions was acquired at a concentration of 50 µM. As well as many published ^{1}H-^{15}N HSQC spectra of Mdm2, triple resonance protein NMR experiments have been performed previously to fully assign Mdm2. These assignments have been published in the BMRB (BMRB ID: 2410[300] and 15945[306]) and were used to partially assign the TROSY spectrum of Mdm2. Ambiguity in the central region precluded the assignment of some peaks and since the His_6-tag was not removed additional peaks corresponding to His_6-tag are present and could not be assigned. However, 76% of identified peaks were able to be assigned based on the previous literature assignments. The partially assigned spectra is shown in **Figure 92**.

Figure 92. The partially assigned ^{1}H-^{15}N TROSY of Mdm2. The peaks that have been assigned are shown.

With the majority of peaks assigned, the binding of ligand **157** can be explored. The p53 binding pocket of Mdm2 is composed of fifteen amino acids (**Table 15**), some or all of which would be expected to shift upon binding of a small molecule to the p53 pocket. However, it is very common for other peaks, located on the periphery of the binding pocket to shift. Examination of the sets of amino acids reported to shift upon binding of a small molecule to Mdm2 differs from study to study and even between members of the same family of inhibitors. For example, a study by McDonnell *et. al.* revealed that for 12 isoindolinone inhibitors with highly conserved structures, different chemical shift changes were observed indicating that slight changes in ligand structure will change the binding mode and result in a different set of amino acids exhibiting a CSP.[306] However, there are a few residues, common across the majority of studies, that are reported to shift upon binding of a small molecule. The amino acids of the three binding pockets that p53 occupies on Mdm2 along with other amino acids that commonly undergo CSP's upon binding of a small molecule are shown in **Table 15**.[265,298,306] All of the residues in **Table 15** have been successfully assigned.

Table 15. The amino acids of the p53 binding pockets on Mdm2 and other important amino acids that commonly shift upon binding of a small molecule.

Important Residues

L26 Pocket	W23 Pocket	F19 Pocket	Others
100	S92	Y67	M50
T101	L54	E69	L57
V53	G58	H73	I61
	Y60	I74	H96
	F91	V75	I99
		M62	T26
		V93	V108
			E52
			F55
			L85
			K94

6.3 The Qualitative Binding of Ligand 157 to Mdm2

This section discusses the binding of ligand **157** (**Figure 93**) to Mdm2 using a ^{1}H-^{15}N TROSY. Ligand 157 in acetone-d_6 was added to 50 μM Mdm2 to a final concentration of 70 μM, in the same solvation conditions as described previously. In total, 10% acetone-d_6 was added to the final solution to improve the solubility of the ligand. A control experiment of the *apo*-protein with 10% acetone-d_6 was performed to confirm that any CSP's were due to ligand binding. The spectra of the *apo*-protein with 10% acetone-d_6 perfectly matched the spectra with no acetone-d_6. Upon adding ligand **157** to Mdm2, significant peak shifts were observed for a number of residues, many of which are located either in the binding pocket or on the periphery **(Figure 94**). The CSP's were calculated as a Euclidean distance, as described by **Equation 12** with a value of $\alpha = 0.14$, calculated from the ratio of the chemical shift range (in ppm) of ^{1}H to that of ^{15}N (3.08/22.1 = 0.139). Deciding which shifts are large enough to be considered an indicator of the binding site is not straightforward. The standard method is to calculate the standard deviation, σ, of the Euclidean CSP and any residues that display a CSP >σ are considered significant.[291] This cut-off is also commonly increased to >2σ to increase the specificity of the analysis and generate less false positives. Those residues with a CSP >2σ are labelled in **Figure 94**.

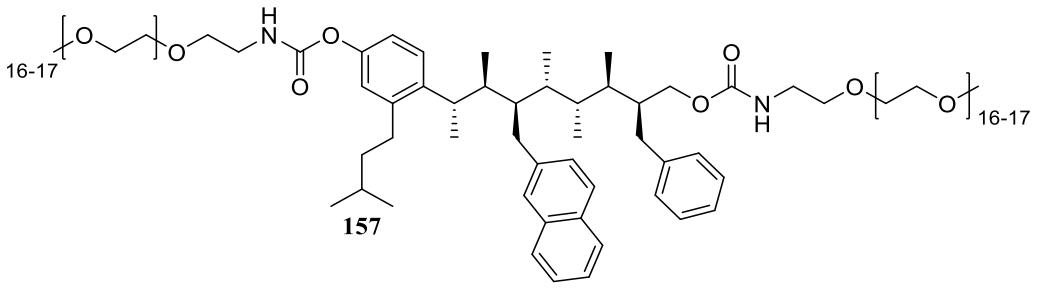


Figure 93. Ligand **157**

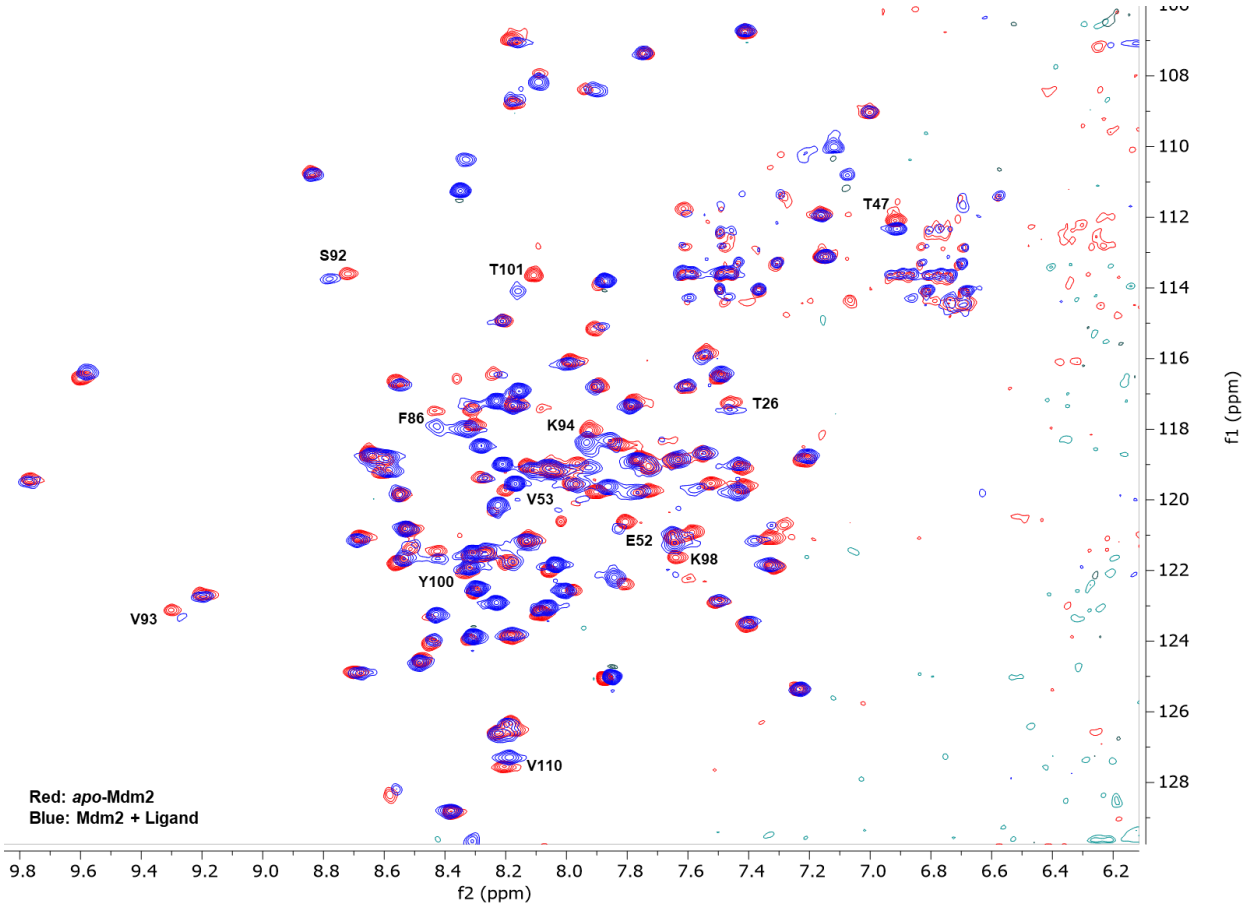


Figure 94. The ^{1}H-^{15}N TROSY of Mdm2 (red) and Mdm2 **157** (blue). The significant CSP's are noted.

The largest chemical shift changes are observed for residues in the Leu26 pocket, with all three residues (Y100, T101, V53) that define the L26 pocket shifting by >2σ. Large chemical shift changes (>2σ) are also observed for residues V93 and S92 which are both located in the F19 pocket. Of the remaining residues that induce a large chemical shift change, F86, K98, K94, T26 and E52 are located on the periphery of the binding pocket and commonly observed to shift upon binding of a small molecule. An additional 23 residues exhibited a moderate CSP (CSP > 1σ) many of which are located in the W23 and F19 pockets. A table summarising the measured CSP's and their calculated Euclidean distance is shown in section 9.7.2 of Chapter 9. The Euclidean distances for all residues with a CSP > 1σ is >0.04 ppm

which is commonly considered the lower threshold that indicates whether that the amino acid is involved in binding to the ligand.[293]

The residues that exhibit CSPs have been mapped onto Mdm2, shown in **Figure 95**. The residues coloured in red are those with a CSP > 2σ, those coloured in orange are those with a CSP > 1σ and those coloured in yellow are those with a CSP < 1σ. It is clear from **Figure 95** that most of the residues that shift are located in or on the periphery of the p53 binding pocket. The most perturbed regions of Mdm2 upon binding of **157** are composed of residues F91 to Y104 and residues M50 to T63. The first set of residues are located on an α-helical domain and β-sheet that forms the sides of the cleft near the F19 pocket whilst the second set of residues are located on a second α-helical domain that forms the side of the cleft that runs along the entire p53 binding region.

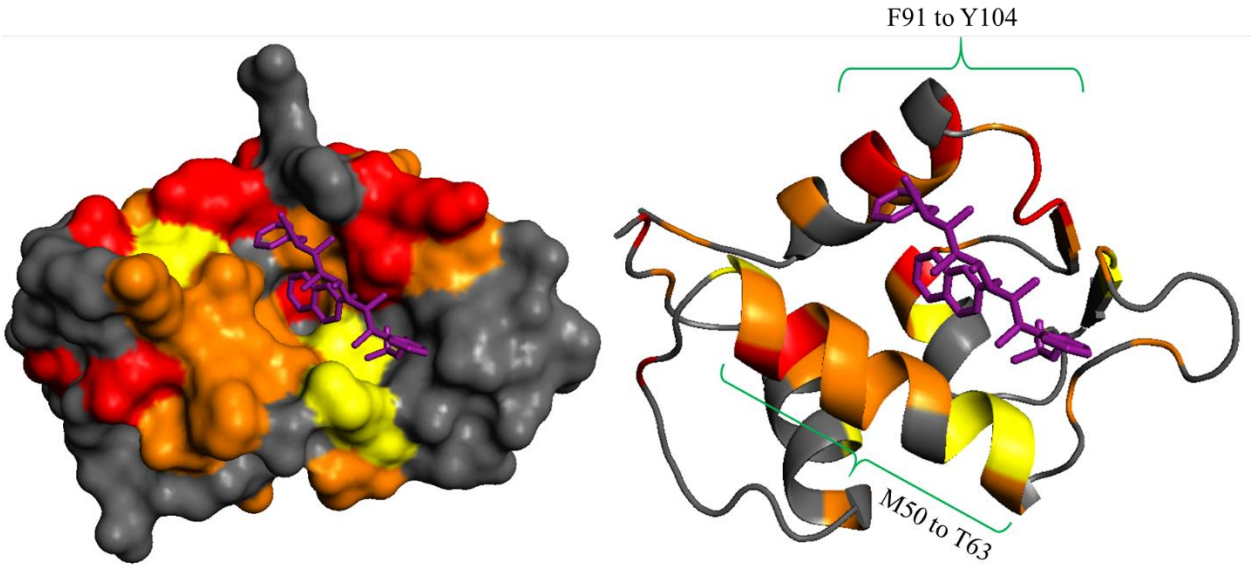


Figure 95. The binding of **157** to Mdm2 deduced by ^{1}H-^{15}N TROSY. Residues coloured in red experience a large CSP (>2σ), residues coloured in orange experience a moderate CSP (>1σ) whilst those coloured in yellow experience a weak CSP (<1σ). PDB ID: 1YCR

CSP's are also observed for a few residues not in the binding pocket (Y47, V28, L33, V110, D68, S40). These changes could be attributed to conformational changes in Mdm2 upon binding of **157**. It is well known that ligands can induce a conformational change in either the whole protein or in regions of its secondary structure to accommodate the ligand. CSP's observed for residues outside of the binding pocket can be caused by this secondary effect rather than direct binding of **157** with this region of Mdm2. This is supported by the fact that the residues not in the p53 binding pocket that have shifted are isolated from one another. If ligand **157** were binding to a second site on Mdm2, one would expect to see a group of chemical shifts of residues located close to one another shift. Reports by Hamilton *et. al.* (V28, D68)[265] and Holak *et. al.* (Y47, V28)[301] report similar residues located outside of the binding pocket exhibiting significant CSP's and attribute them to conformational changes in the protein.

6.4 The Quantitative Binding of Ligand 157 to Mdm2

With the binding of ligand **157** to Mdm2 confirmed, the estimation of the K_d between ligand **157** and Mdm2 was performed using a ligand titration by ^{1}H-^{15}N HSQC spectroscopy, as described previously. Although it would make analysis much simpler using the ^{1}H-^{15}N TROSY due to the enhancement in resolution and sensitivity, it is not possible as this requires 12 hours of spectrometer time per experiment. Therefore, a ^{1}H-^{15}N-SOFAST HMQC, a derivative of the HSQC experiment that uses short inter-scan delays to reduce experimental time, was performed. The SOFAST-HMQC does however come at a loss of sensitivity and resolution, and not all residues are well resolved. A protein concentration of 70 µM was used and five ligand concentrations ranging from 17.5 to 280 µM were used (**Figure 96**). Ligand concentrations >300 µM could not be explored as precipitation of the ligand was observed. Therefore, it is possible that δ_{max} will not be identified and a large error in the estimation of K_d will be obtained. Mbinding (software released by MestreNova) was used to analyse the CSP's and calculate the K_d. The measured CSP's are used to generate a calculated binding curve from which δ_{max} is estimated and K_d is calculated using **Equation 11**. The K_d was calculated for a total of 20 peaks, 18 of which corresponded to assigned peaks, and an average K_d and σ was calculated.

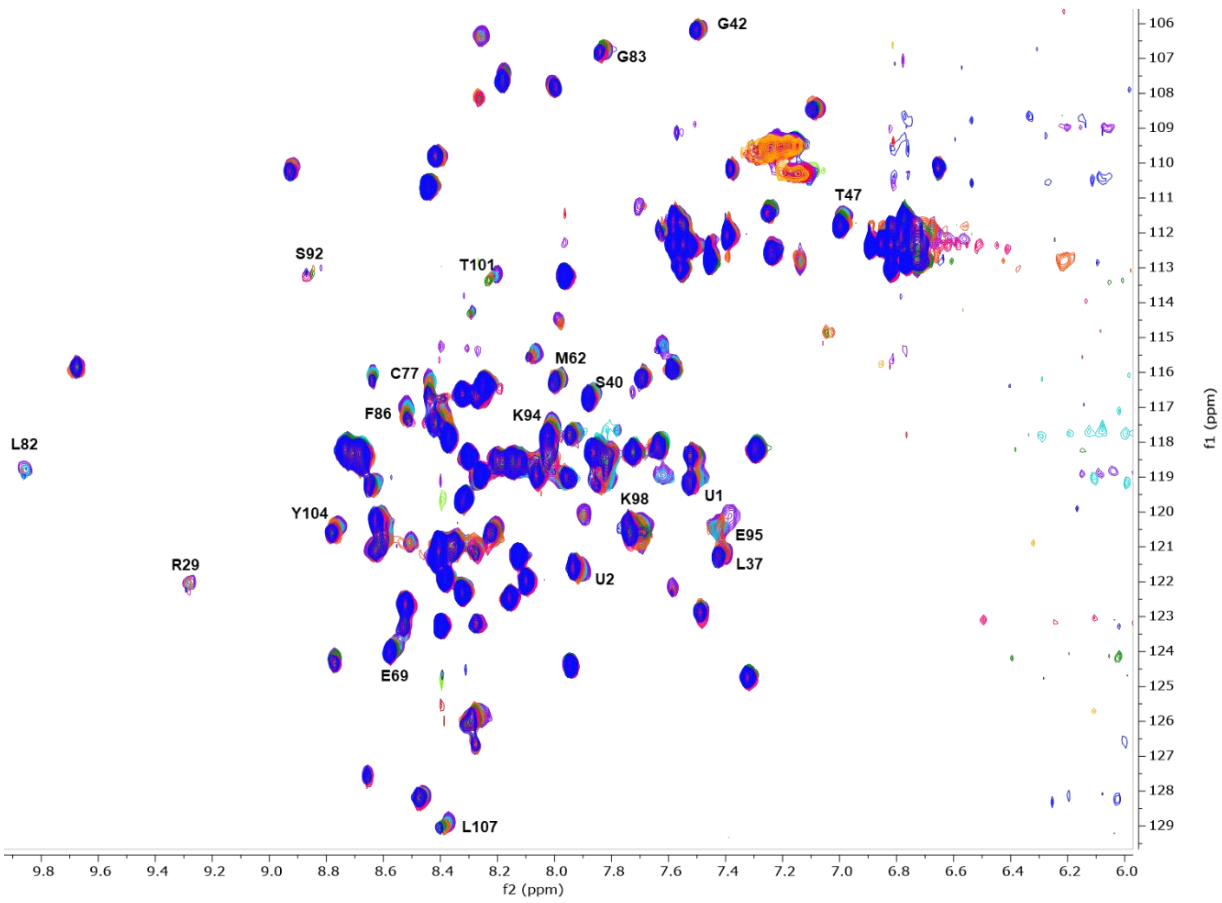


Figure 96. The titration of ligand **157** against Mdm2. Purple (70 µM Mdm2), light blue (70 µM Mdm2 + 17.5 µM ligand), orange (70 µM Mdm2 + 35 µM ligand), green (70 µM Mdm2 + 70 µM ligand), pink (70 µM Mdm2 + 140 µM ligand), dark blue (70 µM Mdm2 + 280 µM ligand). The 20 peaks that were considered in the calculation of K_d are highlighted.

Although the CSP's seem moderate, the low concentrations of ligand required to induce these shifts resulted in an average K_d of 17.2 ± 4.5 μM (**Table 16**). The data shows that for the residues where CSP >2σ, the lower values of K_d were observed, confirming that these residues are interacting strongly with **157**. An exemplar binding curve for T101 and a zoom in of this peak is shown in **Figure 97**.

Table 16. The calculated value of K_d and maximum CSP for 20 different residues interacting with **157**. The average K_d and σ K_d are also shown.

Residue	K_d (μM)	CSP Max (ppm)
C77	16.9	0.09
E95	16.0	0.03
F86	13.2	0.06
G42	20.3	0.02
G83	25.0	0.03
K94	12.3	0.06
K98	12.9	0.05
L107	15.2	0.04
L37	23.4	0.07
L82	21.5	0.04
M62	20.9	0.03
R29	17.1	0.05
S40	18.2	0.03
S92	11.6	0.06
T101	10.5	0.07
T47	12.1	0.05
Y104	15.3	0.04
E69	21.0	0.03
unassigned1 (U1)	24.2	0.05
unassigned2 (U2)	16.2	0.03
Average K_d (μM)		**17.2**
σ K_d (μM)		**4.5**

The value of K_d obtained suggests a moderate affinity between ligand **157** and Mdm2. Although the molecular docking simulations performed by AutoDock Vina suggested that ligand **157** would bind to Mdm2 with only a slightly lower affinity than that of Nutlin-2, experimentally there is a large difference in the strength of binding (K_d = 90 nM).[195] This is likely due to the fact that during the AutoDock simulations, solvation effects are not considered at all. Ligand **157** has much poorer aqueous solubility than the Nutlins, possibly causing the reduction in binding affinity observed experimentally. If the aqueous solubility of **157** is not sufficient then the concentration of **157** in solution will be lower than expected, resulting in a weaker binding affinity calculated. The binding affinity experimentally measured for **157** more closely resembles those reported for the terphenyls synthesised by Hamilton *et. al.* which ranged from 0.18-297 μM depending on the hot-spot groups attached to the terphenyl scaffold.[265] The aqueous solubility of these ligands was also reported to be problematic.

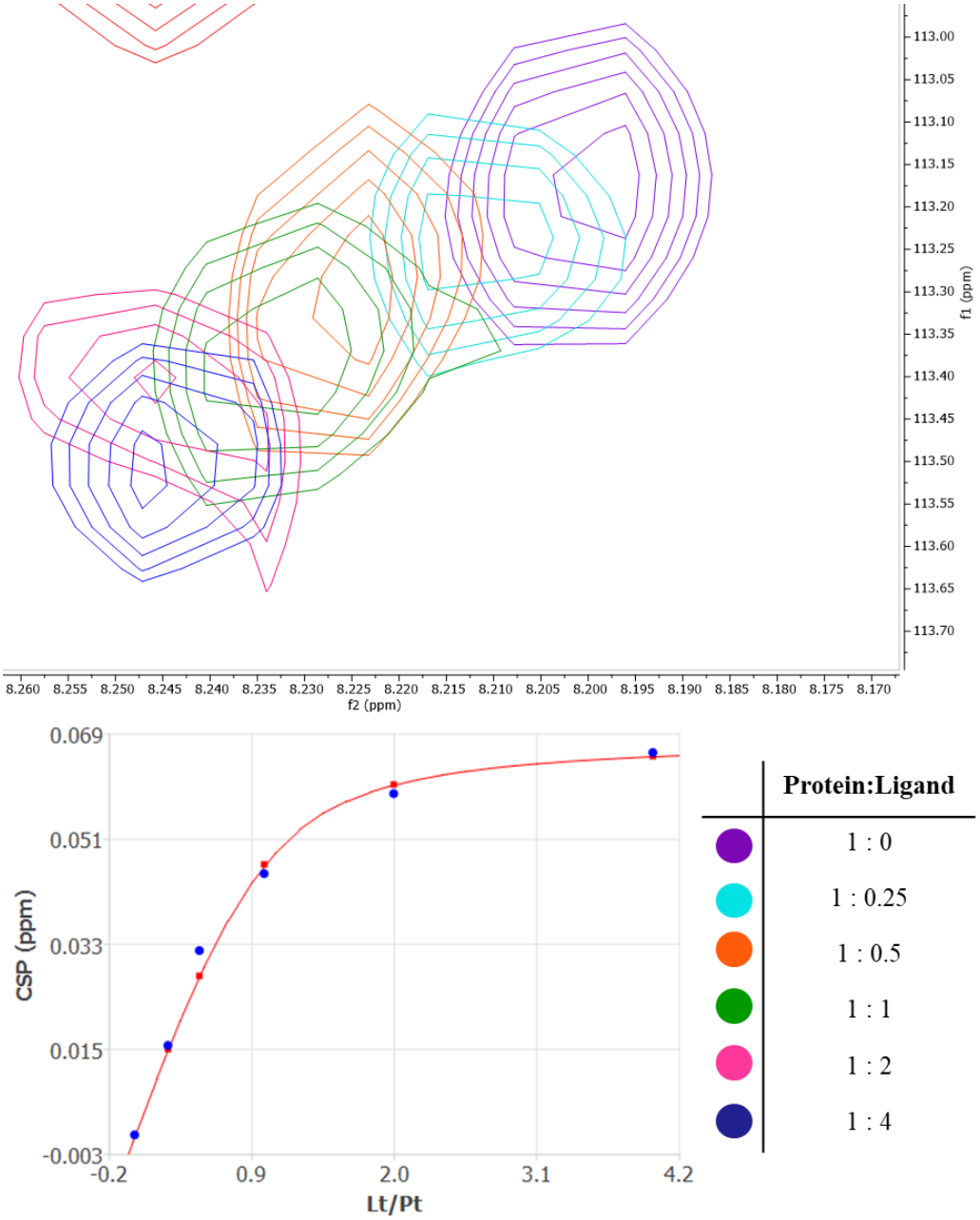


Figure 97. A close up of the CSP for T101 with increasing concentrations of **157**. The experimentally determined CSP (blue) was used to calculate a binding curve (red) from which a value of K_d could be calculated using a nonlinear least-square fitting of Equation 11. The analysis was done using Mbinding from MestreNova.

6.5 The Binding of Ligands 164, 172 and 179 to Mdm2

In addition to ligand **157**, three other ligands (**164**, **172**, and **179**, **Figure 98**) were synthesised and their binding to Mdm2 was explored experimentally using the ^{1}H-^{15}N SOFAST HMQC. A titration of all three ligands at varying concentrations against Mdm2 was performed.

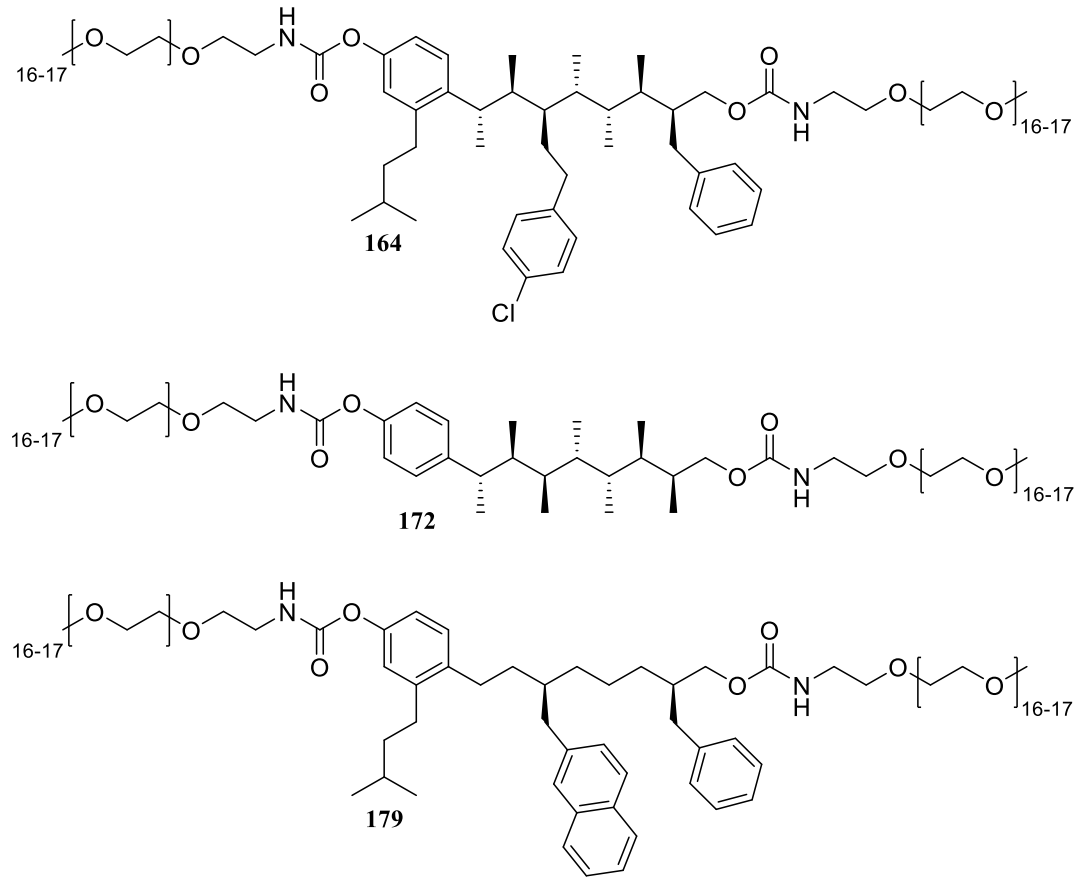


Figure 98. Ligands **164**, **172** and **179**

The first ligand explored was ligand **164** (**Figure 98**), which is expected to display a stronger binding affinity to Mdm2 than that of ligand **157** due to the introduction of the Cl atom that should occupy a deep void in the Trp23 pocket, which has been reported to be critical for competitive binding.[265] The binding of ligand **164** to Mdm2 is shown in **Figure 99**. A protein concentration of 70 μM was used and five ligand concentrations ranging from 17.5 to 280 μM were used. Ligand concentrations >300 μM could not be explored as precipitation of the ligand could be observed. A total of 20 residues, 18 of which have been assigned, were used in the analysis of $K_{d.}$ A very similar set of residues exhibit a CSP upon binding of the two ligands **157** and **164**, however the CSP's observed for ligand **164**, compared to ligand **157**, are larger.

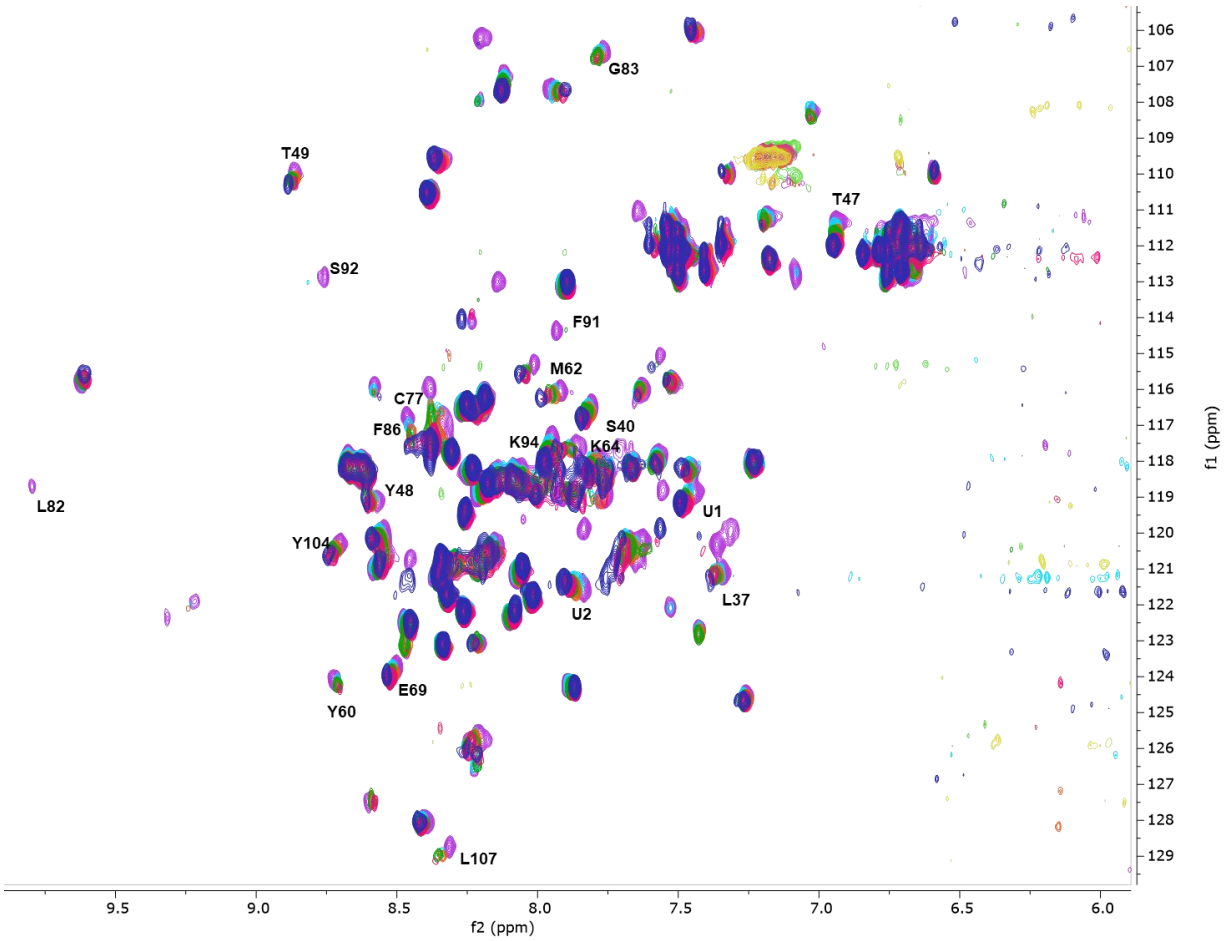


Figure 99. The titration of ligand **164** against Mdm2. Purple (70 μM Mdm2), light blue (70 μM Mdm2 + 17.5 μM ligand), orange (70 μM Mdm2 + 35 μM ligand), green (70 μM Mdm2 + 70 μM ligand), pink (70 μM Mdm2 + 140 μM ligand), dark blue (70 μM Mdm2 + 280 μM ligand). The 20 peaks that were considered in the calculation of K_d are highlighted.

A comparison of the ^{1}H-^{15}N SOFAST HMQC spectra in **Figure 96** and **Figure 99** clearly shows that very similar peaks shift upon binding of **164** and **157**. This suggests that the two ligands are binding in an analogous fashion to Mdm2 and interacting with similar residues. A stronger binding affinity of 8.8 ± 3.1 μM was observed for **164**, confirming the importance of the correct placement of the Cl atom in obtaining an improved binding affinity. The calculated K_d of all 20 peaks are shown in **Table 17**. Although every effort was made to analyse the same set of peaks for **164** and **157**, challenges with spectral overlap or resolution precluded the analysis of some peaks for **164** that were analysed for **157**. Alternative peaks that could be confidently analysed were chosen as an alternative. Despite a few peaks missing, it can clearly be seen, from the comparison of **Table 17** with **Table 16**, that many of the residues that interact strongly with ligand **164** also interact to the greatest degree with ligand **157**.

Table 17. The calculated value of K_d and maximum CSP for 20 different residues interacting with **164**. The average K_d and σK_d are also shown.

Residue	K_d (μM)	CSP Max (ppm)
C77	11.8	0.09
E69	3.3	0.05
F86	5.4	0.13
F91	7.3	0.06
G83	6.9	0.05
K64	9.7	0.06
K94	6.9	0.11
L107	8.0	0.08
L37	12.4	0.05
L82	11.1	0.04
M62	14.5	0.08
S40	10.6	0.07
S92	3.7	0.13
T47	7.4	0.10
T49	10.0	0.06
U1	7.2	0.07
U2	11.7	0.07
Y104	8.6	0.07
Y48	6.7	0.04
Y60	13.1	0.05
Average K_d (μM)		**8.8**
σ K_d (μM)		**3.1**

The effect of the Cl atom can be confirmed by examining residues that are located deep in the W23 pocket, such as L85, which is located on an α-helix that forms the bottom of the binding cleft. It has been suggested by Hamilton *et. al.* that strong CSPs of the L85 residue are indicative of strong binding due to occupancy of the deep void in the W23 pocket.[265] Whilst the naphthyl derivative **157** does induce a small CSP in L85, it is notably larger with the Cl derivative **164** (**Figure 100**), suggesting a much stronger interaction between ligand **164** and L85 than ligand **157** and L85.

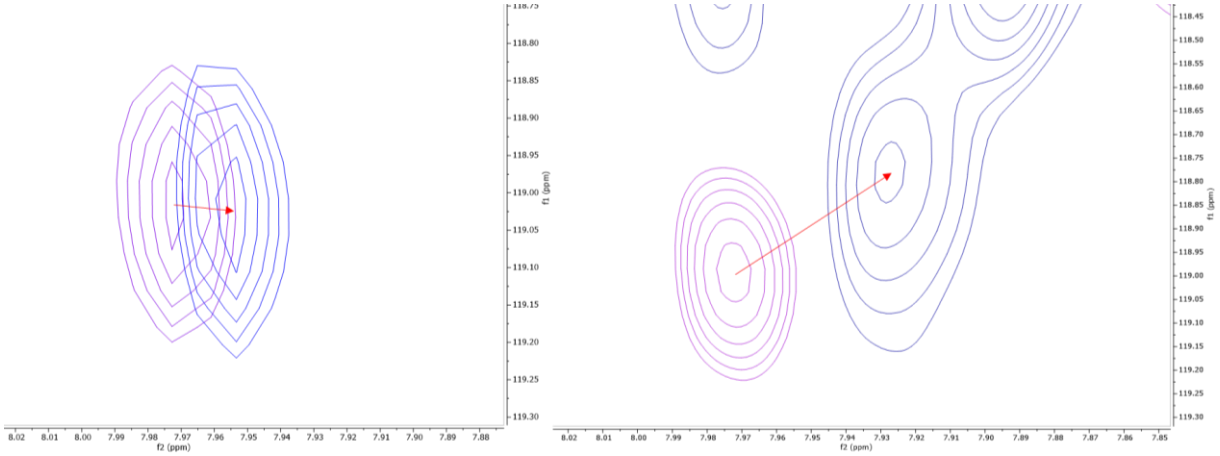


Figure 100. The effect of ligand **157** (left) and ligand **164** (right) on the L85 residue. It can clearly be seen that the CSP induced by **164** for L85 is much more significant, indicating a stronger interaction between **164** and the L85 residue. Purple (70 μM Mdm2), dark blue (70 μM Mdm2 + 280 μM ligand).

As well as L85, G83 and L82 are also present deep in the W23 pocket, both of which exhibit a more significant CSP upon binding of ligand **164** compared to ligand **157**. The results obtained from residues G83, L82 and L85 confirm that ligand **164** is able to penetrate deeper into the W23 pocket than ligand **157**, as a result of the Cl atom occupying the deep void in the W23 pocket, as predicted by molecular docking (**Figure 101**).

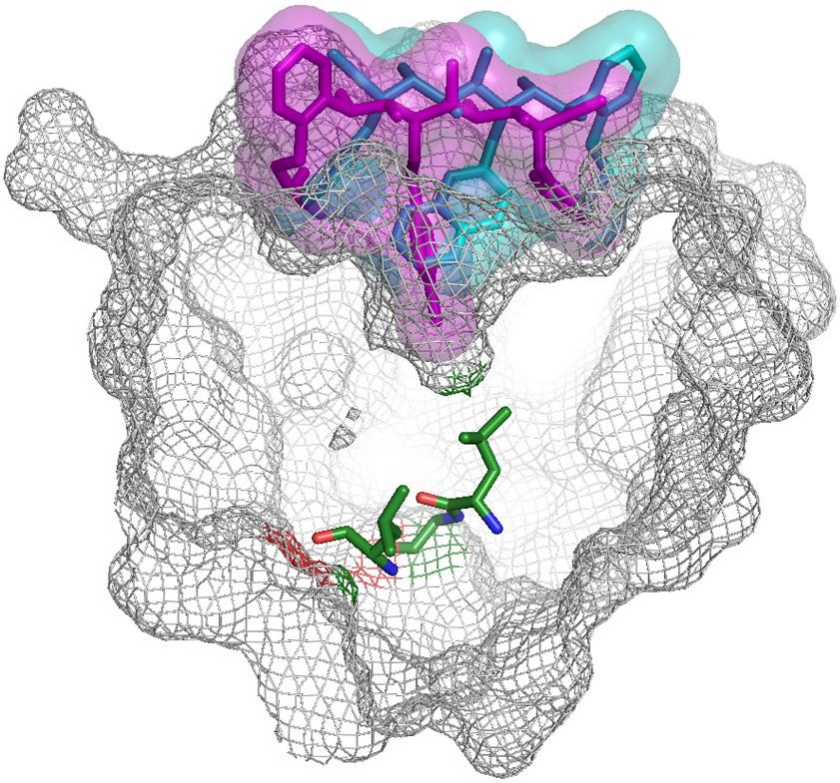


Figure 101. The binding of ligand **157** (blue) and ligand **164** (purple) as predicted by AutoDock Vina. Residues L82, G83 and L85 are shown in green.

In addition to the two ligands **157** and **164**, which display moderate binding to Mdm2, two control ligands were synthesised. Ligand **172** was synthesised to confirm the importance of the hot-spot groups in binding to Mdm2 whilst ligand **179** was synthesised to confirm the importance of preorganising the free conformation of the ligand for improved binding affinities. Both ligand **172** and **179** are expected to bind weakly, if at all, to Mdm2. The binding of both ligands was explored using the ^{1}H-^{15}N SOFAST HMQC.

The binding of ligand **172** revealed very few significant CSPs, with the chemical shift of many residues remaining unchanged on increasing concentrations of ligand **172** (**Figure 102**). A protein concentration of 70 μM was used and five ligand concentrations ranging from 35 to 350 μM were used. The lack of CSPs observed confirms that ligand **172** binds poorly to Mdm2.

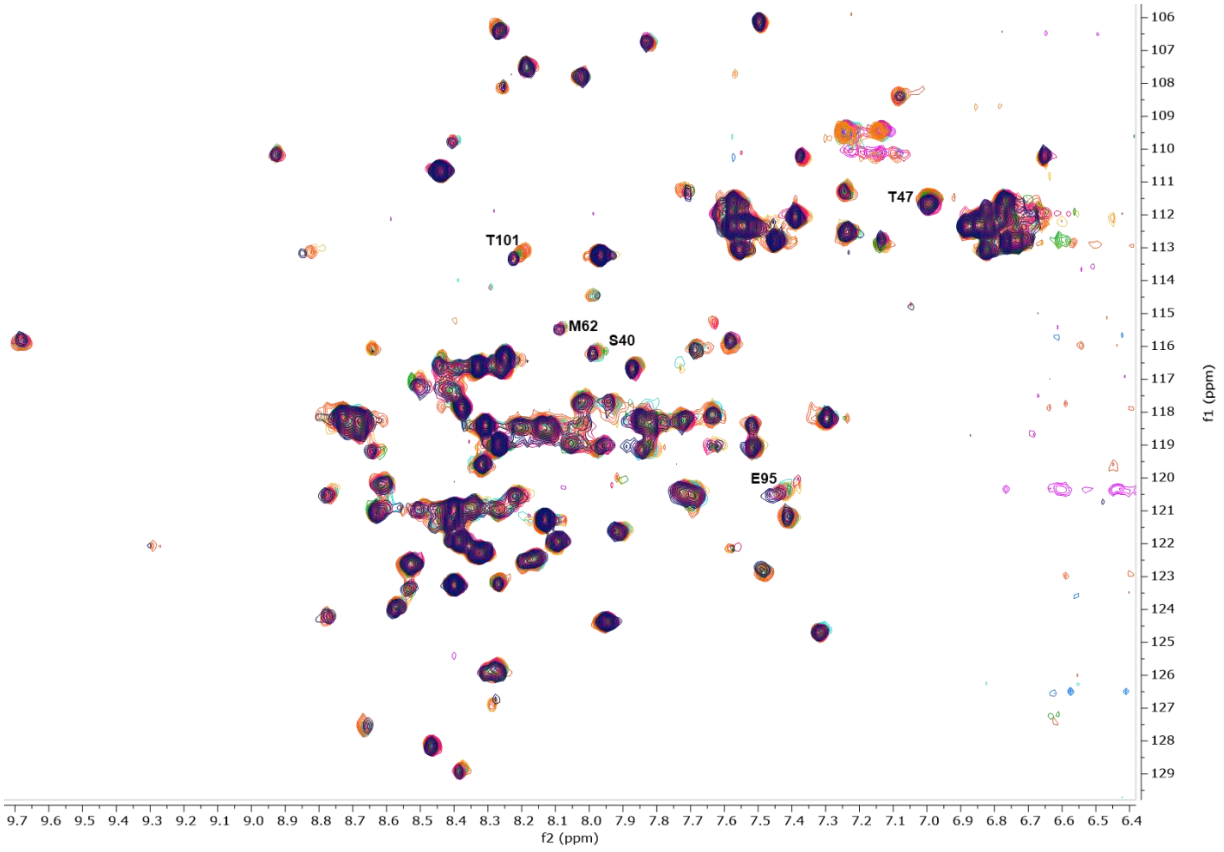


Figure 102. The titration of ligand **172** against Mdm2. Purple (70 μM Mdm2), light blue (70 μM Mdm2 + 35 μM ligand), orange (70 μM Mdm2 + 70 μM ligand), green (70 μM Mdm2 + 140 μM ligand), pink (70 μM Mdm2 + 280 μM ligand), dark blue (70 μM Mdm2 + 350 μM). The 5 peaks that were considered in the calculation of K_d are highlighted.

Due to the lack of CSPs, the quantitative analysis of binding of ligand **172** is much more challenging. Furthermore, it is more likely that δ_{max} will not be attainable and a large error will be associated with the calculated value of δ_{max} and K_d. Indeed, it is clear from the binding curve of M62, shown in **Figure 103**, that δ_{max} is not achieved and a poor binding curve is obtained. Therefore, we cannot reliably measure K_d for the control molecule **172.** However, it is clear that **172** is binding very weakly to Mdm2, confirming the importance of the hot-spot groups in obtaining strong inhibition of Mdm2.

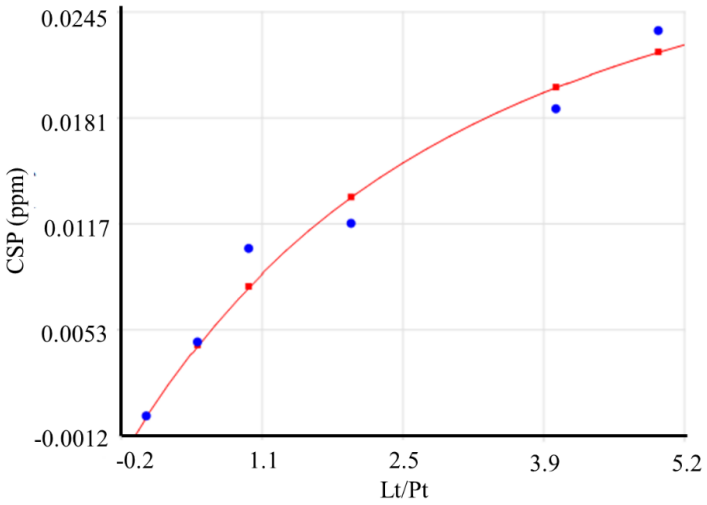


Figure 103. The binding curve for M62 from the binding of **172** to Mdm2. The red line is a non-linear least square fitting to **Equation 11**

The binding of the second control ligand **179** was explored, to confirm the importance of conformational preorganisation. A protein concentration of 70 μM was used and six increasing concentrations of ligand **179** were used ranging from 35 μM to 350 μM. The binding of ligand **179** to Mdm2 revealed a small set of residues, many of which are located in the binding pocket of Mdm2, exhibiting small CSP's upon increasing concentrations of **179** (**Figure 104**). However, these CSP's were much smaller in magnitude when compared to ligands **164** and **157** and required larger concentrations of **179** to induce these CSP's. In analogy to control ligand **172**, poor binding curves were obtained for control molecule **179** because δ_{max} could not be experimentally attained. Therefore, a reliable value of K_d was not obtained. However, the lack of measurable CSPs confirms that the control molecule **179** is binding weakly to Mdm2, compared to the two ligand **164** and **157**. This result confirms that conformationally preorganising the free conformation of a ligand can improve its binding affinity.

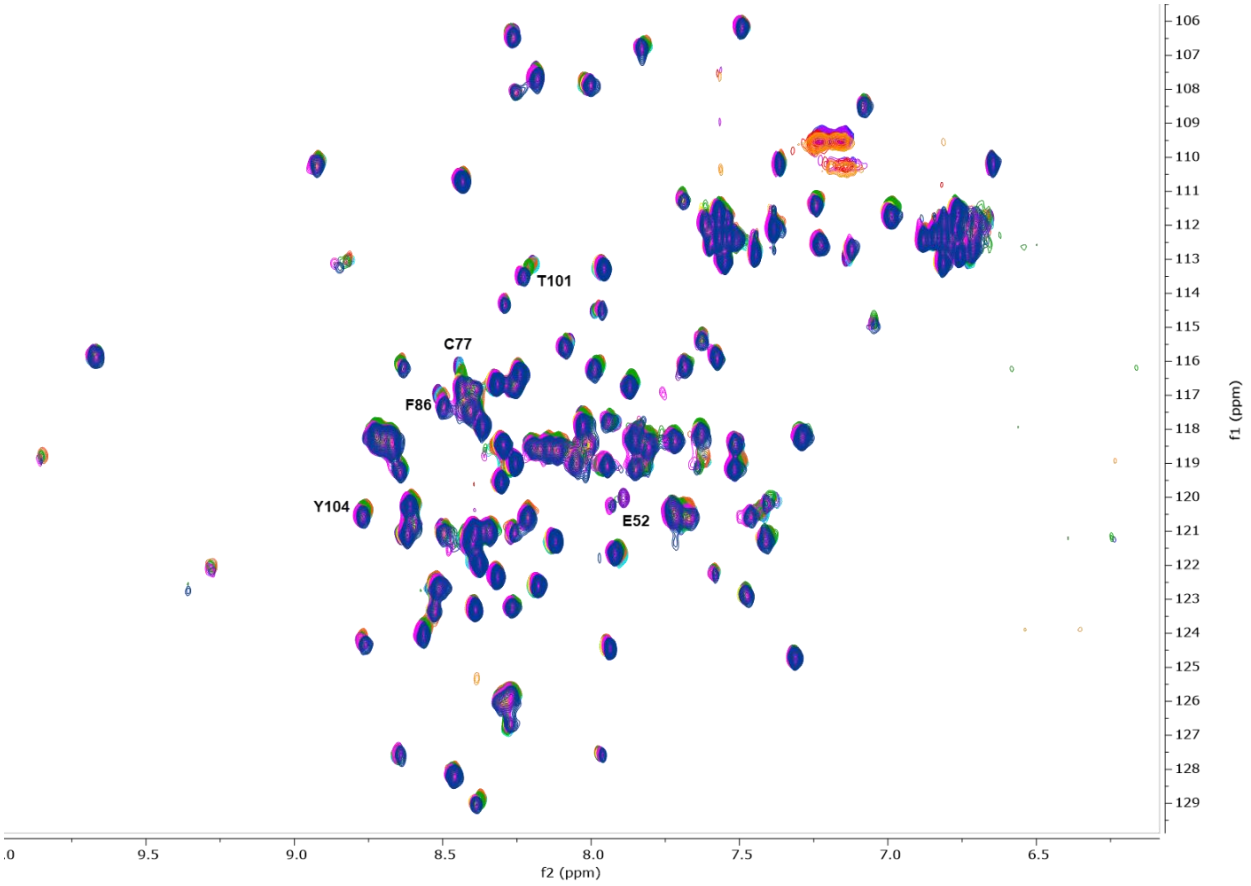


Figure 104. The titration of ligand **179** against Mdm2. Purple (70 μM Mdm2), light blue (70 μM Mdm2 + 35 μM ligand), orange (70 μM Mdm2 + 70 μM ligand), green (70 μM Mdm2 + 105 μM ligand), yellow (70 μM Mdm2 + 140 μM ligand) pink (70 μM Mdm2 + 280 μM ligand), dark blue (70 μM Mdm2 + 350 μM). The 5 peaks that were considered in the calculation of K_d are highlighted.

6.5 A Summary of the Binding of Ligands 157, 164, 172 and 179 by NMR spectroscopy

The CSP's observed in the ^{1}H-^{15}N SOFAST HMQC or TROSY spectra of ligands **157**, **164**, **172** and **179** confirms that the p53 mimetics **164** and **157** bind to the p53 cleft on Mdm2. This was deduced from the observation of CSP's for many residues located in the p53 binding pocket whilst only a few CSP's were observed for residues located far from the p53 pocket, which are likely the result of conformational changes upon binding of a ligand. In contrast, the control ligands **172** and **179** do not induce significant CSP's and those residues that do shift upon binding of the control ligands do not shift to the same magnitude as seen for ligands **157** and **164**. Additionally, greater concentrations of the control ligands are required to induce the shifts. This suggests that the control ligands **172** and **179** have a lower affinity to Mdm2 than ligands **157** and **164**, confirming the importance of both the hot-spot groups and conformational preorganisation in obtaining higher affinity therapeutics.

For the two designed p53 mimetics (ligand **157** and ligand **164**) the CSP's were tracked with increasing concentrations of the ligands and using Mbinding from MestreNova the binding curves were calculated and a value of K_d was estimated. A summary of the experimentally determined values of K_d are shown in **Table 18**.

Table 18. A summary of the experimentally determined K_d values from ^{1}H-^{15}N-SOFAST HMQC spectroscopy for ligands **157**, **164**, **172** and **179**

Ligand	K_d (μM)
157	**17.18**
164	**8.81**

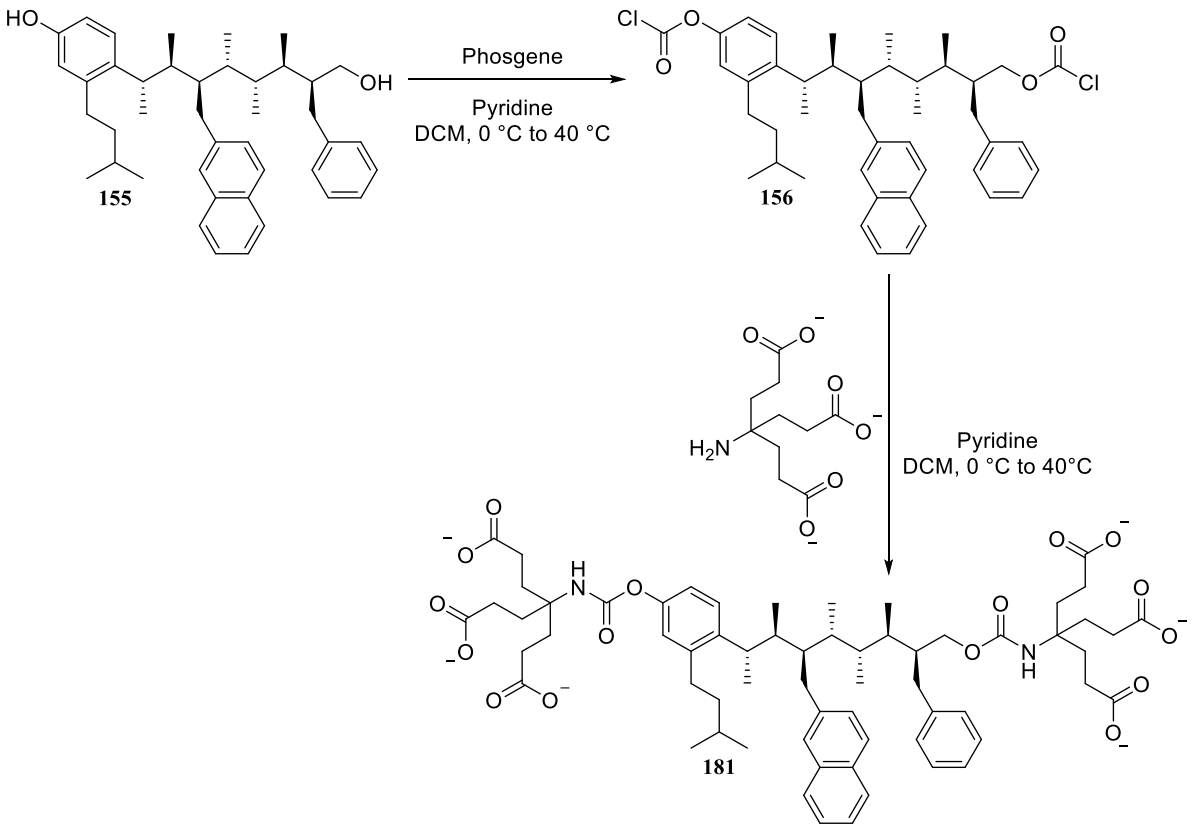
Ideally, the binding of the ligands would be explored by different biophysical methods to confirm the results obtained from ^{1}H-^{15}N SOFAST HMQC spectroscopy, however the poor solubility of the ligands prevents the use of certain methods due to their sensitivity. Since the aqueous solubility of the ligands is far from ideal, alterations to the scaffold to improve the water solubility should be performed before a wide range of biophysical methods are tested. Alternative biophysical methods that could be used following improvements made to the solubility of the ligands is discussed in future work.

Chapter 7: Future Work

7.1 Improving the Solubility of the p53 Mimetics

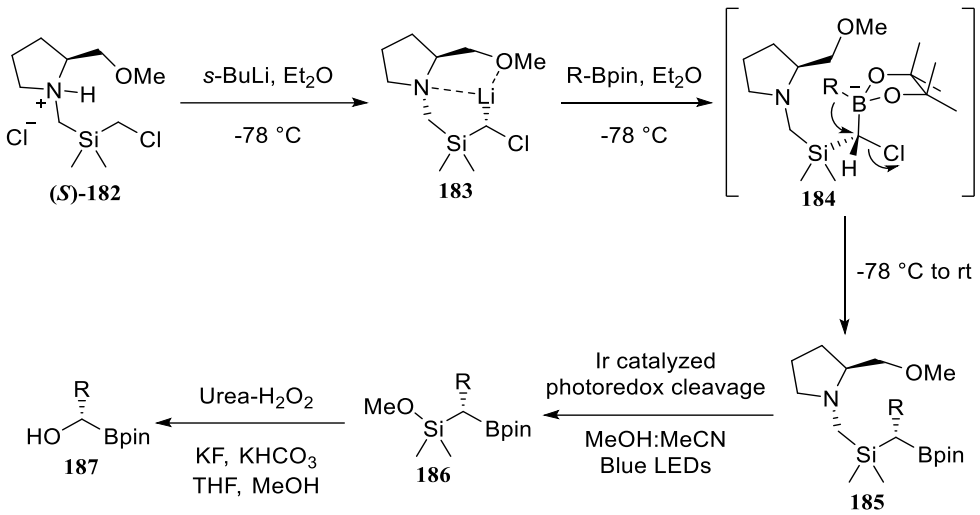
Although the p53 mimetics designed and synthesised demonstrate reasonable affinity to Mdm2, the pharmacokinetic properties of the molecule are far from ideal, restricting their application as potential therapeutics. The lipophilicity of the molecules should be addressed before their biological activity is fully explored as currently their poor aqueous solubility will lead to large uncertainties in any binding constants measured experimentally. Additionally, the poor solubility will result in some analytical methods being unusable due to high concentrations of ligand relative to protein required. With a calculated LogP (calculated using ACD labs) of 8.97 the big-PEGylated p53 mimetic (**157**) is far from the ideal LogP of 5, governed by Lipinski's rules.

One strategy to improve the LogP is to add better solubilising groups onto the two terminal positions of the scaffold, in place of the PEG chains. One example that has proven successful in the design of water-soluble glucose receptors is the addition of multicharged water solubilising dendrimer chains. The dendrimer chains contain multiple anionic groups and the synthetic glucose receptors are readily soluble in aqueous solvents.[309,310] The impressive water solubilising properties of dendrimers has been exploited previously in drug design, with examples of dendritic therapeutics having been reported in the literature.[311,312] One promising application of dendrimers is their use in drug conjugation. It has been demonstrated that a dendrimer unit can be conjugated to a therapeutic molecule via a cleavable linker such as disulphide bridges.[313] Using this method, it is possible to ensure that the dendrimers do not interact with the receptor but enable hydrophobic drugs to be delivered to their site of action. A future endeavour of this project could be to explore the addition of multianionic dendrimer chains to the scaffold. A possible synthetic route is shown in **Scheme 26**. A very similar synthetic sequence for installation of the dendrimer chains to the installation of the PEG chains can be proposed with a change in the amine used. The amine needed can be synthesised from the reduction of nitromethanetrispropionic acid, which is commercially available.



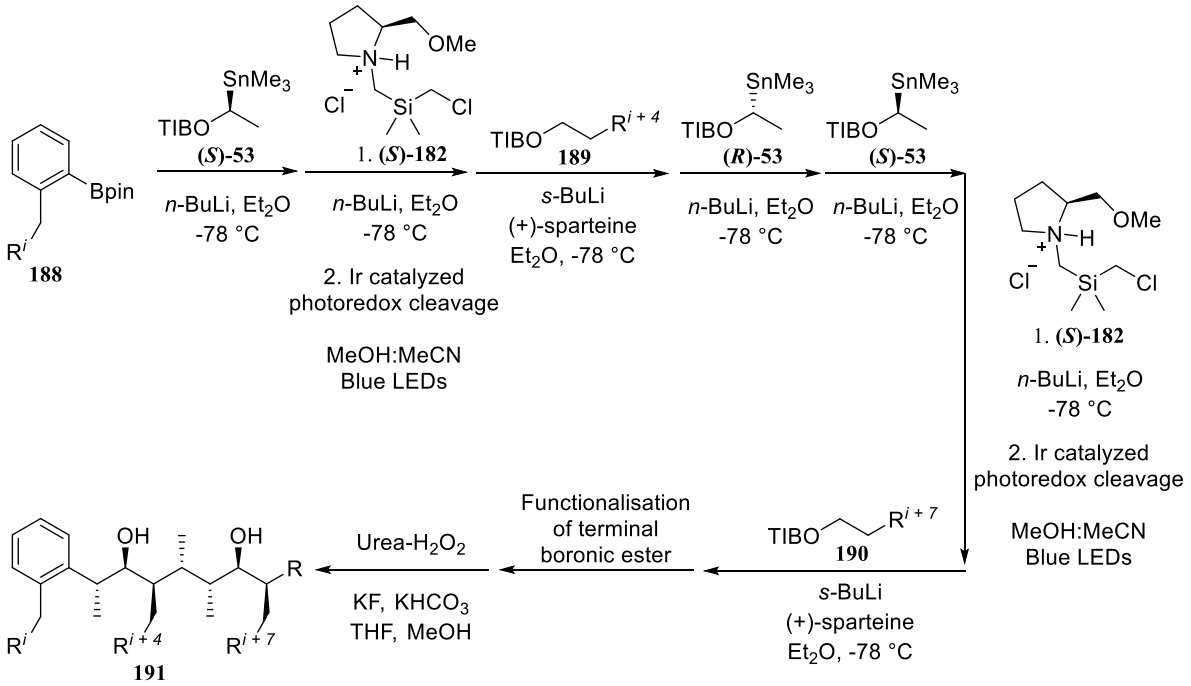
Scheme 26. The proposed synthesis of a p53 mimetic functionalised with dendrimers, **181**.

The addition of solubilising groups at the terminal positions requires no redesign on the lithiation–borylation sequence used to synthesise the mimetics themselves. However, adding large polar groups at either end of long nonpolar molecule could encourage the formation of large aggregates. Instead, it should be explored whether it is possible to incorporate solubility onto the backbone of the molecule instead, or in addition to the terminal positions. Recently Bootwitcha *et. al*. have reported an iterative assembly line process towards the synthesis of polypropionates through the introduction of carbinol units.[258] The polypropionate motif features alternating methyl and hydroxy groups along a hydrocarbon backbone. The synthesis of this motif is extremely challenging often requiring multiple redox processes. Bootwitcha *et. al.* have demonstrated that enantioenriched lithiated α-chlorosilanes, **183**, can be used in the reagent-controlled homologation of boronic ester to generate α-silylalkyl boronates, **186** (**Scheme 27**). Oxidation of the silyl group reveals the corresponding hydroxy group, **187**. The resulting boronic ester can either be subjected to further homologations with either enantioenriched lithiated α-chlorosilanes to install contiguous hydroxy groups or α-lithiated benzoate esters to install side chains such as methyl, as previously described.



Scheme 27. The proposed mechanism of the stereocontrolled homologation of lithiated α-chloromethyl silane with a boronic ester.

Using this methodology, hydroxy groups could be incorporated onto the backbone of the scaffold, in place of the methyl groups, such as structure **191** (**Scheme 28**). The hydroxy groups could be left intact or further functionalised to introduce either hot-spot groups or solubilising groups. A potential pitfall of this approach is that the *syn*-pentane control, previously enforced by the methyl groups, is no longer as dominating as the hydroxy group is considerably smaller in size than the methyl group. The effect of introducing hydroxy groups onto the scaffold would need to be fully investigated using MM conformational search calculations and QM calculations prior to synthesis.



Scheme 28. A proposed synthesis of a p53 mimetic bearing hydroxy groups on the backbone, **191**, using the homologation of boronic esters.

The final option for introducing greater solubility into the molecule is to change the hot-spot groups added onto the scaffold, to introduce more polar groups. For example, p53 features a tryptophan residue as one of the hot-spot groups. The NH of the indole of tryptophan is engaged in a hydrogen bond with a leucine residue on p53. Incorporation of an indole group onto the scaffold would not only improve solubility slightly but also potentially increase the binding affinity of the mimetic with Mdm2, although it has been suggested that this hydrogen bond is not necessary for competitive binding. The introduction of the indole group using lithiation–borylation is challenging as the NH of the indole will need to be protected. The choice of protecting group will need to be fully explored to ensure that lithiation on the indole does not occur. For example, the BOC protecting group, which is a common choice of amine protecting group, would likely direct lithiation onto the indole. A potential protecting group that could be used is the benzyl group. Future work would be to explore the installation of protected amine groups onto the scaffold using lithiation–borylation.

7.2 Further Experiments to Investigate Binding of our Designed Mimetic to Mdm2

Chapter 6, discussed the use of ^{1}H-^{15}N HSQC spectroscopy to analyse the binding of the designed mimetics to Mdm2. This methodology allowed us to confirm the location of binding of the mimetics and provided an estimate of the K_d. However, other biophysical methods should also be explored in future work to further explore the binding of the designed mimetics to Mdm2.[314]

Previously, the use of protein-based NMR methods were used to analyse protein-ligand binding. Protein-based methods are advantageous as they can be used to directly monitor changes to protein conformation and structure upon ligand binding.[287] It can provide atomistic information about protein-ligand binding, such as which amino acids are interacting with the ligand. Additionally, protein-based methods have been reliably used to determine dissociation constants (K_d).[283] However, they typically require considerable amounts of isotopically enriched protein and acquisition of 2D-NMR data can be very time demanding. Furthermore, protein-based methods are typically only possible for smaller proteins of <50 kDa.[314] An alternative to protein-based methods are ligand-based NMR methods whereby changes to the spectrum of the ligand are monitored upon binding to the protein. Using ligand-based methods, the binding of a ligand to a protein can be studied using unlabelled proteins of any size.[282] The most popular ligand-based NMR method is saturation-transfer difference (STD) spectroscopy.

During an STD experiment, the protein is irradiated at a specific radiofrequency field that selectively targets the protein resonances.[315] This is known as the *on-resonance* spectrum. The magnetisation is transferred from the protein to any ligands that are bound to it, by spin diffusion, which will be transferred to ligands free in solution since the bound and free ligands are in constant exchange.[290] A

second spectrum is acquired using a radiofrequency pulse where no saturation will take place, known as the *off-resonance* spectrum. The difference between the two spectra (the *on-resonance* and the *off-resonance*) is recorded, revealing a spectrum of the ligand featuring only resonances that were saturated by the protein. Protons on the ligand that interact strongly with the protein with therefore receive a greater amount of saturation and will have enhanced NMR signals compared to protons that do not interact with the protein. Therefore, STD can be used to identify the regions of the ligand strongly interacting with the protein. The STD method is very useful is several ligands are to be screened for binding to the protein, as only those ligands that interact with the protein will be observed in the difference spectrum (**Figure 105**).[287] Using the STD method, a large excess of the ligand to protein is typically required, which is challenging for highly lipophilic molecules. Despite this limitation, the STD NMR experiment would be a relatively straightforward experiment to perform in future work and may provide information regarding the regions of the ligand that are strongly interacting with Mdm2. This information could help to improve the design of stronger binding ligands.

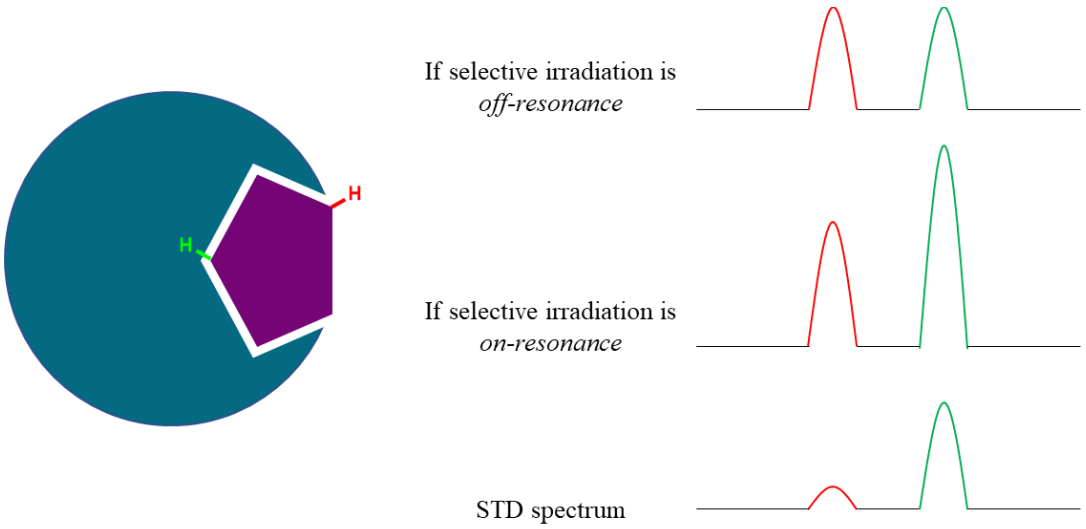


Figure 105. A schematic to demonstrate the principal of saturation transfer difference (STD) NMR experiments to monitor protein-ligand interactions

Protein-ligand interactions are also routinely studied by fluorescence polarisation (FP). The significant advantage of FP assays is that they are amenable to high throughput screening, making them a favourable choice in drug discovery.[316] Additionally, it does not require isotopically labelled proteins. Fluorescence polarisation, also commonly referred to as fluorescence anisotropy, monitors changes in fluorescence which is proportional to the rotation time of a molecule, as a fluorescent ligand is either bound to a protein or competitively displaced (**Figure 106**). If the fluorophore is attached to the ligand itself, then a titration of ligand against protein can be performed and K_d can be measured. If however, a fluorophore is not present on the molecule it is possible to measure K_i through the competitive displacement of a known inhibitor that possess a fluorophore. As the ligand displaces the known inhibitor a loss is anisotropy is expected. If this method is chosen, the competitive inhibitor needs to

bind to the protein with a similar affinity to the designed ligand so a strong idea of the K_d between the ligand and protein is required.

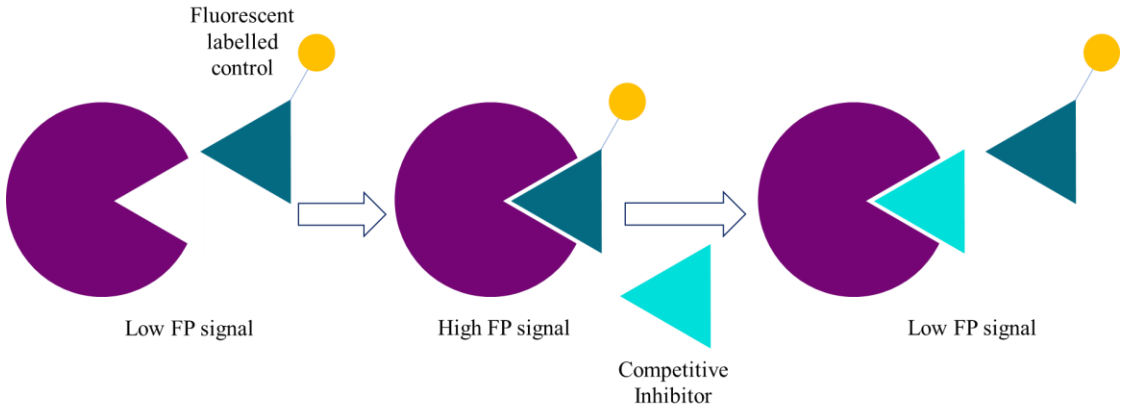


Figure 106. A general schematic to show the competitive fluorescence polarisation assay.

Since a fluorophore is not present on our designed p53 mimetics, the more suitable approach would be to bind a competitive inhibitor with a fluorophore to Mdm2 and displace this with the designed p53 ligands, to obtain a value of K_d. Preliminary work using this method has been conducted by Dr. Madhavachary Rudrakshula however, the assay was unsuccessful and inconclusive results were obtained. A fluorescently labelled p53 analogue (Ac-SQETFSDLWKLLPENNVC(Flu)-NH_2) was chosen as the competitive inhibitor however upon binding the peptide to Mdm2, a K_d of 1.86 ± 0.16 μM was obtained. This binding affinity is lower than those predicted by NMR for the designed p53 ligands and as a result a high concentration of the ligand would be required for competitive displacement. Indeed, at a concentration of 1 mM, poor and scattered competition results are obtained (**Figure 107**). This could possibly be due to precipitation of the ligand at such high concentrations resulting in a much lower effective concentration of ligand in solution. Future work would be to design a p53 peptide analogue with a lower affinity to Mdm2. This could possibly be done by alanine mutagenesis.

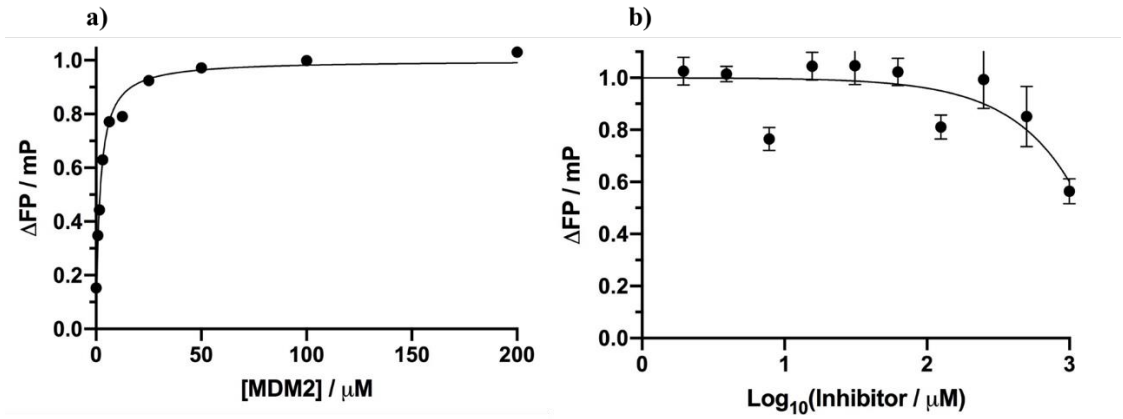


Figure 107. a) The results of the FP assay for the p53 analogue. **b)** The attempt to displace the p53 analogue with ligand **157** was unsuccessful using an FP competition assay.

This project has assumed that the designed p53 mimetics bind to Mdm2 as a linear conformation, thereby reducing the conformational reorganizational energy that is required upon binding as the free ligand conformation is biased towards a linear conformation. Although H-^{15}N HSQC spectroscopy confirmed that the molecules bound to the p53 binding site on Mdm2, it did not provide any detailed information regarding the orientation of binding and the bound conformation of the molecules. In order to explore the bound conformation of the designed p53 mimetics to Mdm2, X-ray crystallography needs to be performed. Not only will X-ray crystallography confirm whether the free conformation of the ligands reflects the bound conformation of the ligands, but it will also allow for visualisation of the protein-ligand interactions and allow for structural optimisation of the hot-spot groups to maximise the binding affinity.[317] However, a major limitation of X-ray crystallography is that it requires suitable crystals of the protein-ligand complex to be grown.

The previous methods discussed do not allow for a direct measurement of the thermodynamic parameters involved with the protein-ligand binding event. One method that can allow for the measurement of such parameters is isothermal titration calorimetry (ITC). ITC works by measuring the heat either released or absorbed during a titration of the ligand into a sample of the protein. The quantity of the heat measured is directly proportional to the extent of binding. ITC can be used to measure the association constant (K_a), the enthalpy changes upon binding (ΔH), from which the Gibbs free energy change (ΔG) and entropy change (ΔS) can be derived. Although ITC allows for a direct measurement of the association constant and the thermodynamic parameters of protein-ligand binding, it typically requires high protein concentrations as it suffers from low sensitivity.[318] In turn this causes considerable challenges in the solubility limits of the titrated compound, with highly aqueous soluble compounds often required. This will be a major limitation in the analysis of the conformationally controlled hydrocarbons using ITC.

It is also possible to use mass spectroscopy (MS) to characterise the protein-ligand complex. Using native-MS, whereby protein-ligand complexes are analysed in their native biological state, it is possible to measure the exact mass of the protein-ligand complex. This confirms that the protein-ligand complex is present and can be used to determine the stoichiometry of binding.[319] Due to the high sensitivity of MS, very small amounts of protein are required. Additionally, the protein does not need to be labelled unlike in NMR. Native-MS however provides no information on the location of binding, and so does not confirm that the ligand has bound in a specific manner to the protein. An additional caveat of native-MS is the more challenging sample preparation which requires the sample to be desalted. The protein-ligand complex needs to remain stable in the desalted conditions for it to be successful.[314] However, the relative ease of acquisition makes native-MS an attractive biophysical technique to test in future work towards confirming the binding of the designed p53 mimetics to Mdm2.

Although many biophysical methods are available for studying protein-ligand interactions, all of them rely on good aqueous solubility of the ligand and a solid understanding of the concentration of ligand in solution. Many of the biophysical methods described may yield inconclusive results due to ambiguity of the concentration of ligand in solution due to possible precipitation. Therefore, before a large range of biophysical methods are screened, the aqueous solubility of the ligand should be addressed first.

7.3 The extension to a 'universal' α-helix mimetic

This thesis has explored the design of a conformationally controlled hydrocarbon scaffold capable of mimicking the *i*, *i* + 3/4 and the *i* + 7 positions of the α-helix. However, as mentioned in section 1.6.2 of Chapter 1, scaffolds that can mimic multiple faces of the α-helix are of high importance as roughly 40% of PPIs mediated by an α-helix display hot-spot groups across two or more faces.[205] The linear scaffold reported in this thesis has 4 distinct binding faces and a promising avenue of future work would be to explore the mimicry of multiple faces of the α-helix with the conformationally controlled hydrocarbon scaffold (**Figure 108**).

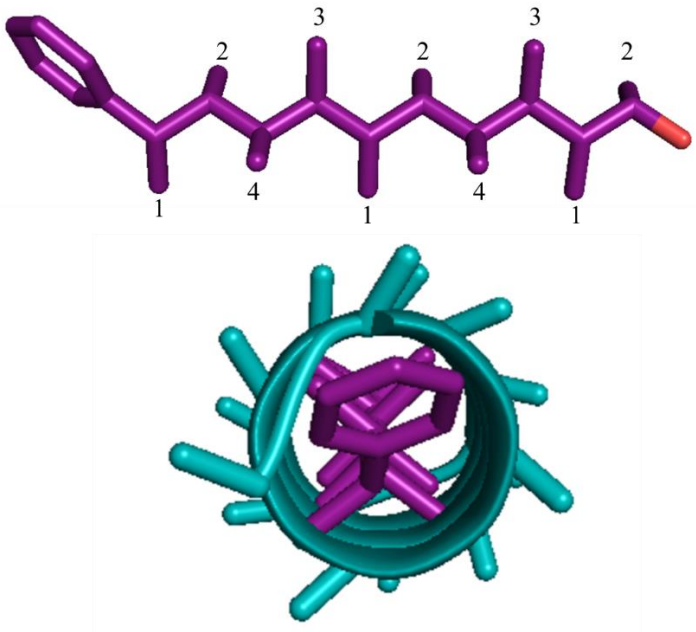


Figure 108. An illustration of the four distinct faces of the linear molecule

We hypothesise that the linear scaffold can be substituted with suitable side chains, using iterative lithiation–borylation, to mimic a highly diverse combination of positions of the α-helix. By simply overlaying the linear scaffold (**11**) with the helix hoping to be mimicked it is possible to identify the positions of the linear scaffold that can mimic the required α-helical residues. Then a simple adjustment of the lithiation–borylation sequence to incorporate appropriate benzoates at judicious positions, will provide access to a conformationally controlled α-helix mimetic of your choice. If successful, we will have access to a universal α-helix mimetic, whereby using the same synthetic methodology and simply

adjusting the reagents used in the sequence, it is possible to mimic any position on the α-helix as desired.[320] Examples of combinations of residues that could potentially be mimicked by the linear scaffold are shown in (**Figure 109**). For example, the first scaffold (**192**) is proposed to mimic the *i*, *i* + 2 and *i* + 5 positions of the α-helix by substituting the 2nd, 3rd and 7th positions on the linear scaffold. Other combinations than the three shown in **Figure 109** are possible.

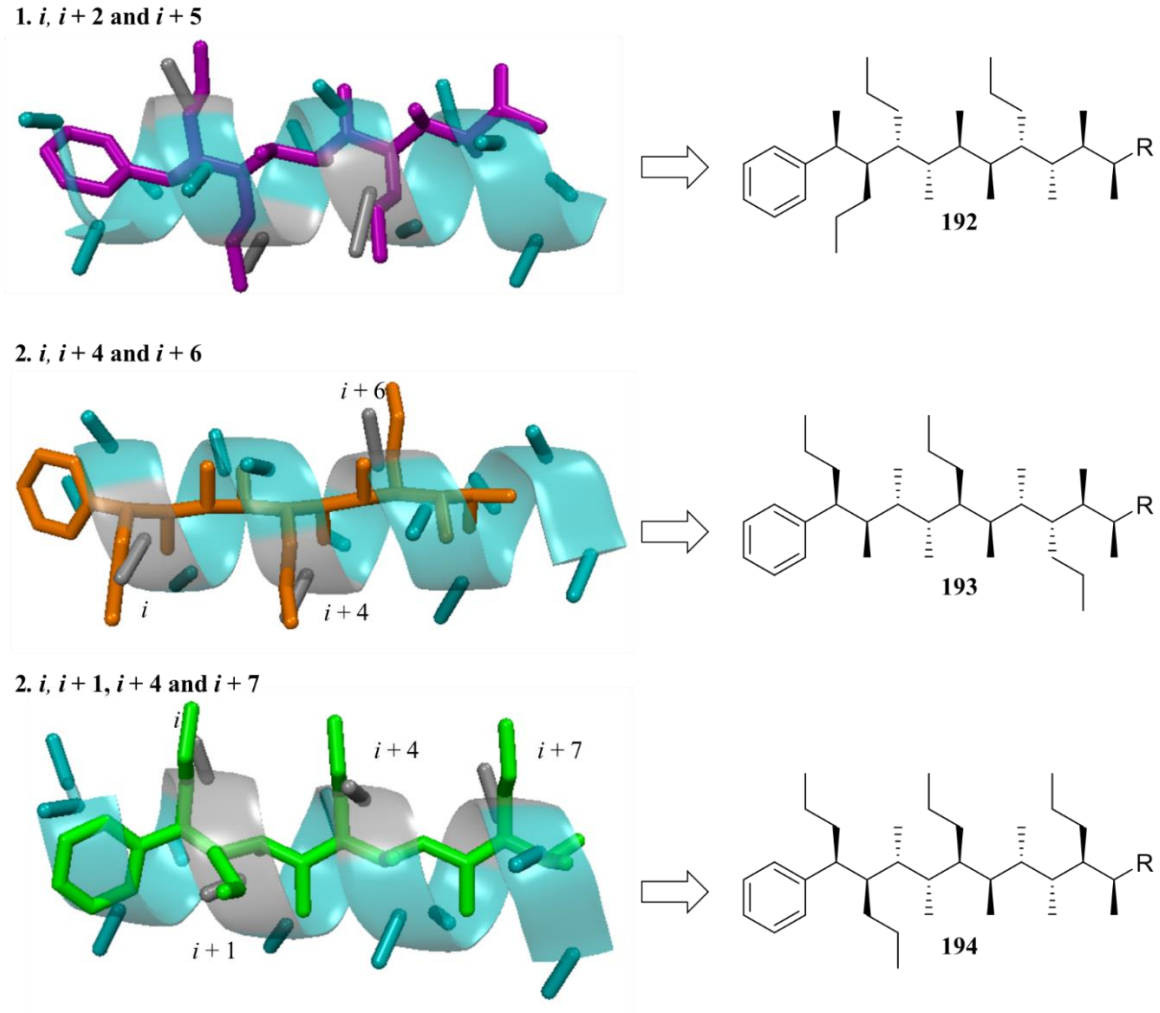


Figure 109. An example of three different sets of residues that could be inhibited using conformationally controlled hydrocarbons. 1 – the mimicry of the *i*, *i* + 2 and *i* + 5 residues of the α-helix requires side chains at the 2nd, 3rd and 7th positions of the linear scaffold (**192**). 2 – the mimicry of the *i*, *i* + 4 and *i* + 6 residues of the α-helix requires side chains at the 1st, 5th and 8th positions of the linear scaffold (**193**). 3 – the mimicry of the *i*, *i* + 1, *i* + 4 and *i* + 7 residues of the α-helix requires side chains at the 1st, 2nd, 5th and 9th positions of the linear scaffold (**195**).

The Mcl1-NoxaB PPI is a key therapeutic target due to its role in regulating apoptosis. Attempts to inhibit this PPI have been reported, and although some success has been met, with a candidate from AstraZeneca now in clinical trials[10], it remains a significant therapeutic challenge. This is because the PPI features a long α-helical region with hotspot groups located across multiple faces. Woolfson *et. al.* recently reported the hot-spot residues of Mcl1-NoxaB, found by computational alanine scanning (**Figure 110**).[321] They found that a central cluster of five residues on NoxaB (Leu11, Arg12, Ile14,

Asp16 and Val18), known as the short motif, and three residues in an extended motif (Leu4, Leu25 and Asn26) were important for binding to Mcl1. Mimicry of the extended motif with a small molecule is a huge challenge, however mimicry of the short motif may be possible, especially considering that the two dominating hot-spot groups are found on the short motif.

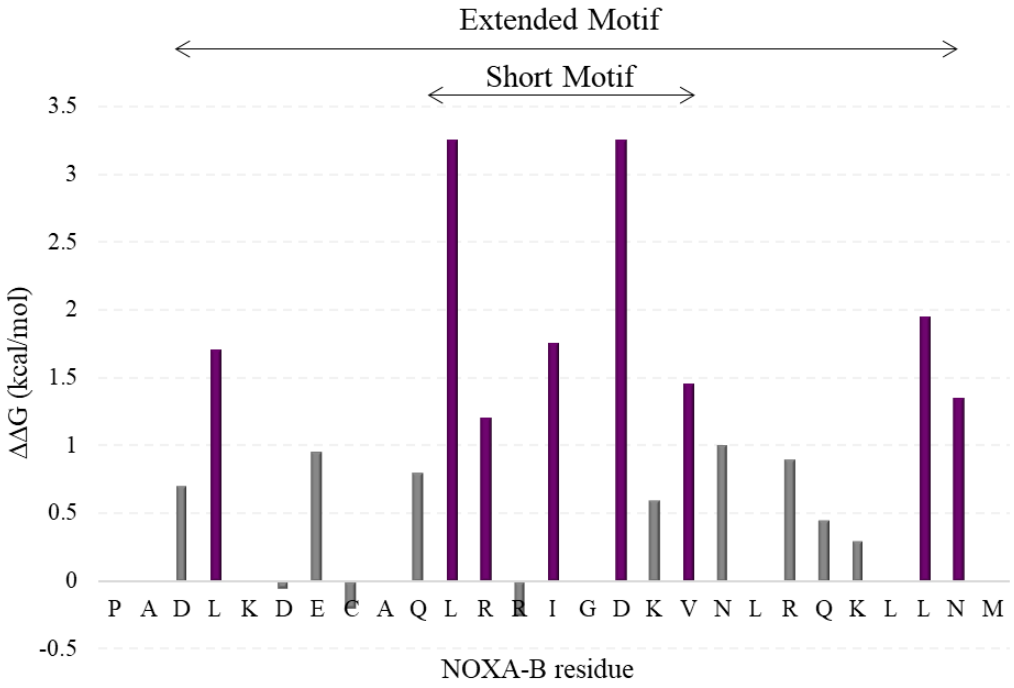


Figure 110. The hot-spot residues of Noxa-B found using computational alanine scanning by Woolfson *et. al.* Residues with a ΔΔG of >4.2 kJ mol^{-1} are coloured in purple. This figure has been reproduced from reference[321].

With hot-spot groups located at the *i*, *i* + 1, *i* + 3, *i* + 5 and *i* + 7 positions, the Mcl1-NoxaB PPI is a good candidate to explore whether conformationally controlled hydrocarbons are capable of mimicking multiple faces of the α-helix. Overlaying the linear scaffold with NoxaB, it is possible to identify the positions that need to be functionalised if mimicry of the *i*, *i* + 1, *i* + 3, *i* + 5 and *i* + 7 positions is to be achieved. The 1st, 3rd, 5th, 7th and 9th position overlay well with these residues. Extending the methyl groups to propyl side chains reveals a good overlap between the 1st, 3rd, 5th, 7th and 9th positions of the linear hydrocarbon and the *i*, *i* + 1, *i* + 3, *i* + 5 and *i* + 7 residues of Noxa-B (**Figure 111**).

The first stage of this project would be to explore the effect on the conformational bias of the linear scaffold following the addition of a larger number of side chains, positioned closer together, when compared to the scaffolds explored in this PhD. Following this, the synthesis of the scaffold would need to be developed as incorporating a mimetic of an aspartic acid residue and an arginine residue is challenging due to the acidic/basic nature of their side chains. Finally, exploration of the conformation of the scaffold by NMR spectroscopy and QM calculations followed by experimentally measuring binding to Mcl1 would need to be explored. If successful, this project would confirm that the

conformationally controlled hydrocarbons, synthesised through iterative lithiation–borylation, show promise as a new class of universal α-helix mimetics.

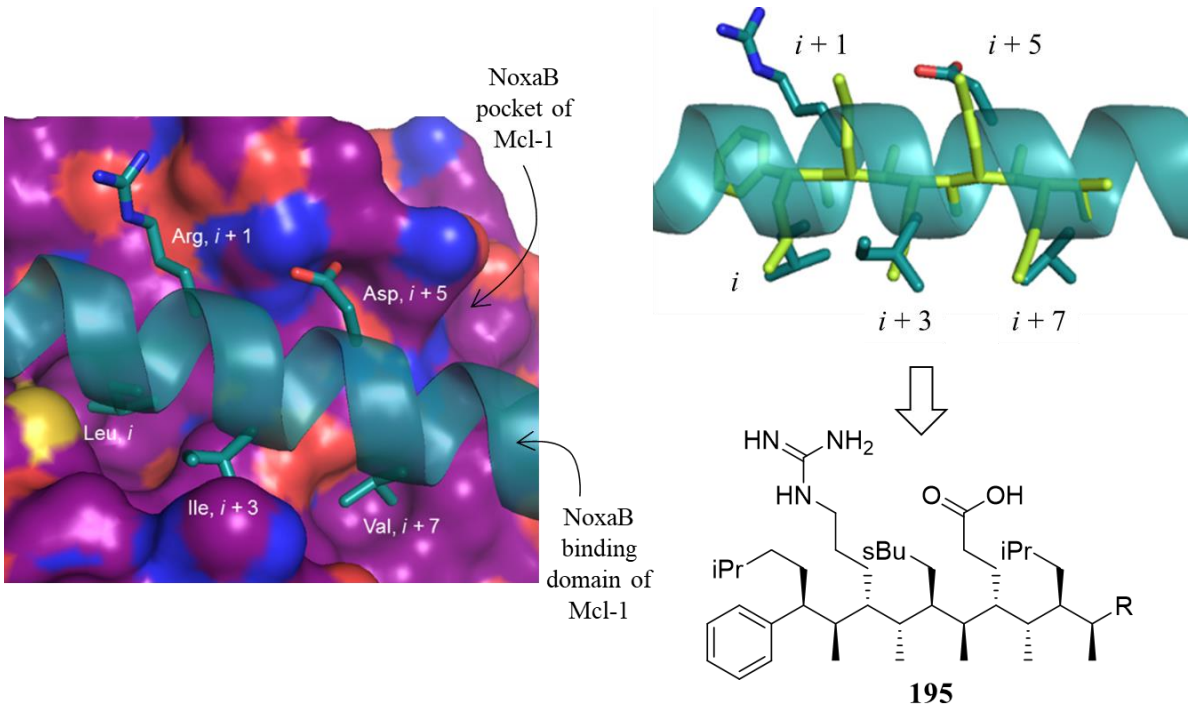


Figure 111. Noxa-B has hot-spot residues at the *i*, *i* + 1, *i* + 3, *i* + 5 and *i* + 7 which could potentially be mimicked by side chains added to the 1st, 3rd, 5th, 7th and 9th positions of the linear scaffold (**195**). A potential mimetic of Noxa-B is shown. PDB ID: 2ROD

Chapter 8: Conclusions

This thesis has demonstrated that contiguously methyl substituted hydrocarbons, conformationally biased due to the avoidance of destabilising *syn*-pentane interactions, are suitable mimetics of the α-helix. The methyl groups positioned at the 1st, 5th and 9th position of the hydrocarbon scaffold **11** closely match the key distances and angular relationships observed between C_β of *i*, *i* + 3/4 and *i* + 7 residues of an α-helix, with an RMSD of 0.85 Å (**Figure 112**).

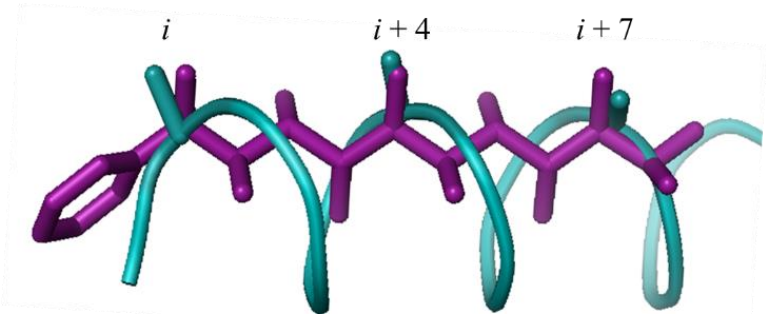


Figure 112. The overlay of a linear hydrocarbon scaffold (**11**) with a model α-helix. It can clearly be seen that the 1st, 5th and 9th methyl groups closely match with the *i*, *i* + 3/4 and *i* + 7 residues.

Small molecule α-helix mimetics are important targets in the inhibition of aberrant PPIs. The binding affinity between proteins engaged in a PPI is typically dominated by a few key residues, known as hot-spot residues, so the search for inhibitors of PPIs has focussed on the mimicry of the hot-spot residues as opposed to the entire protein–protein interface. These hot-spot residues are most commonly found on one face of the α-helix, typically the *i*, *i* + 3/4 and *i* + 7 residues. This thesis explored whether our contiguously methyl substituted scaffold could be substituted with suitable hot-spot groups to mimic the Mdm2 binding domain of p53 and potentially disrupt the p53-Mdm2 PPI. Using lithiation–borylation, the substitution of the methyl group for a suitable mimetic of p53 is as simple as switching the benzoate ester used in the sequence. Using lithiation–borylation it is possible to incorporate mimetics of most amino acid side chains onto the scaffold at any position by designing the correct sequence of benzoate esters to be used.

It was first confirmed that the conformational bias previously observed for the fully methylated hydrocarbon (**11**) was not lost upon replacement of methyl groups for suitable p53 mimetic groups. MM conformational search calculations confirmed that the methyl groups could be directly substituted for propyl side chains (**101**) with no significant loss in conformational control observed with over 73% of conformers adopting a linear conformation. This result was initially unexpected, as the extended side chains have now introduced two *syn*-pentane interactions. However, on closer inspection of the lowest energy conformer found from MM calculations, it can be seen that the dihedral angles have distorted away from their ideal values to accommodate these *syn*-pentane interactions. By introducing an offset

of ~40° between the eclipsing groups in a *syn*-pentane interaction the energy penalty of this interaction will be reduced, resulting in a linear conformational bias still being observed. With a general scaffold successfully designed, the design of a p53 mimetic was explored. The scaffold was initially shortened to reduce the number of homologations required, by incorporating the Leu26 mimetic on the starting aromatic group. The two remaining propyl side chains were transformed to introduce groups to mimic the Trp23 and Phe19 of p53 (**96**). A MM conformational search confirmed that these changes did not perturb the conformational bias, with over 74% adopting a linear conformation (**Figure 113**).

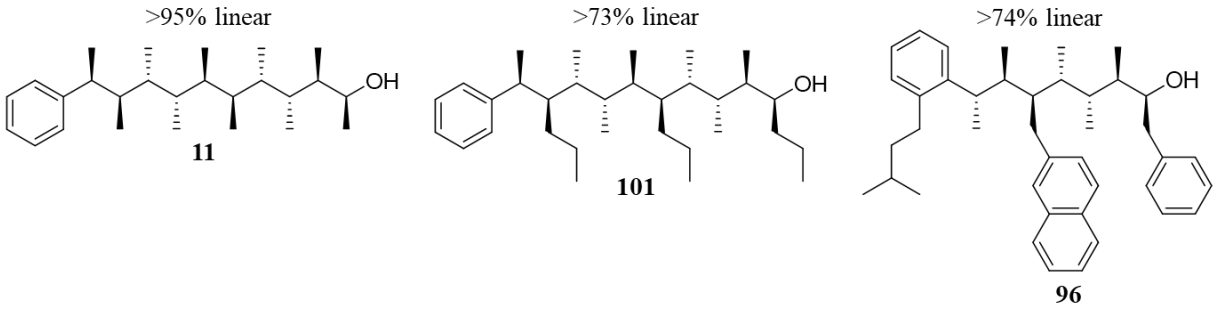


Figure 113. A summary of the key findings from the MM conformational search calculations. The results show that adding longer side chains, or large aromatic hot-spot groups does not perturb the conformational bias, with a linear preference still observed.

The lowest energy conformer of **96** overlays well with p53. Using the pair fitting function in PyMol, and fitting the corresponding carbon atoms of the three 'hot-spot' side chains of **96** with the C_α and C_β atoms of the 'hot-spot' residues on p53 revealed an RMSD of 0.71 Å (**Figure 114**).

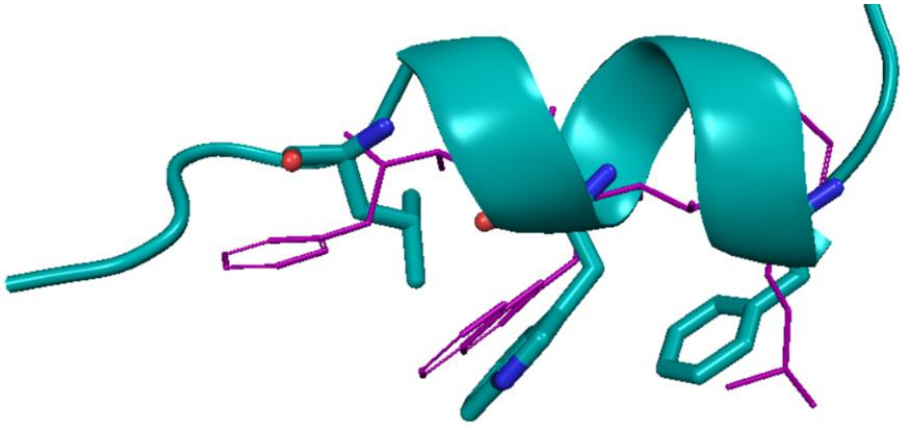


Figure 114. An overlay of p53 (teal) and the lowest energy conformer of the designed p53 mimetic **96** (purple).

Molecular Docking, using AutoDock Vina, confirmed that the designed p53 mimetic (**96**) could bind to Mdm2. A good binding affinity was predicted (-36.8 kJ mol^{-1} vs -41.4 kJ mol^{-1} for Nutlin-2) with the three hydrophobic groups in the three pockets on Mdm2 that are typically occupied by p53. The ligand has bound to Mdm2 in the opposite direction as originally intended, with the Leu mimetic in the Phe19 pocket and *vice versa*. This has been observed by others and has been attributed to a favourable π-π

interaction between the phenyl group on the scaffold and the His96 residue in the Leu26 pocket. Molecular Docking also revealed a deep void in the Trp23 pocket that is not occupied by the naphthyl group. This deep void has been filled previously in the literature by a chlorine atom. Molecular Docking confirmed that by changing the naphthyl group (**96**) to a *para*-chloro phenyl group (**106**) this deep void can be occupied by the chlorine atom and the binding affinity is increased to -38.9 kJ mol^{-1} (**Figure 115**).

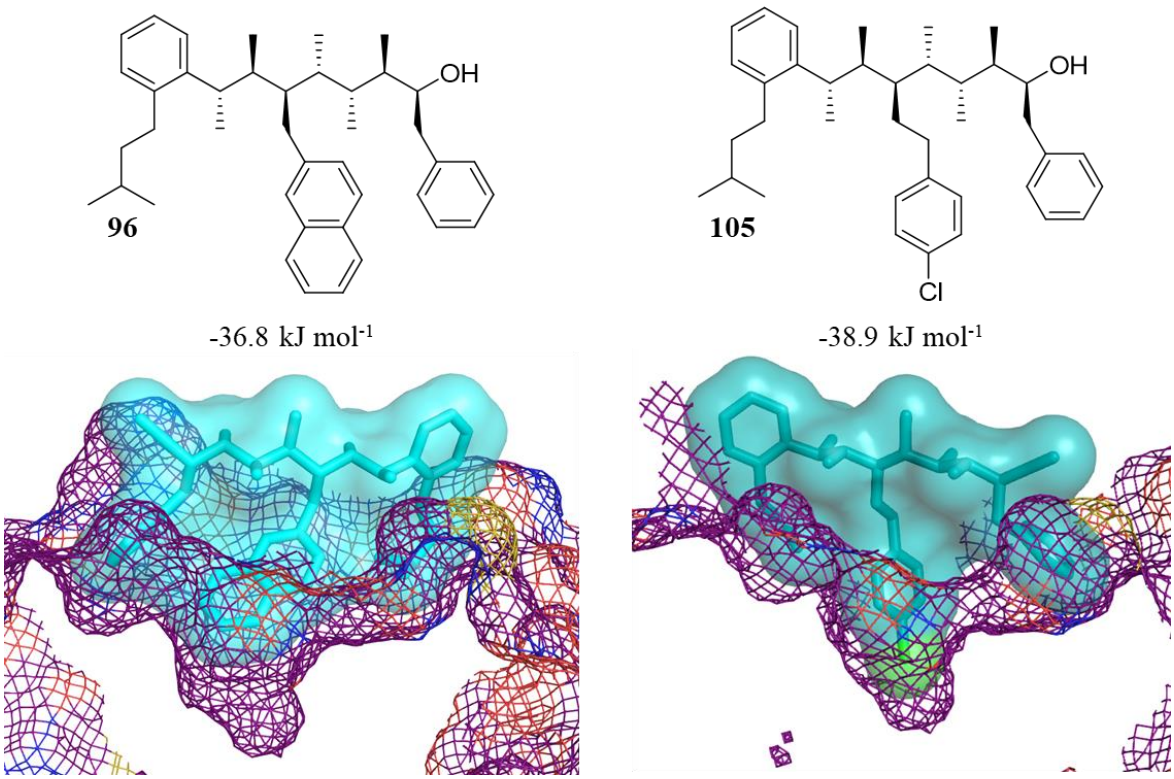
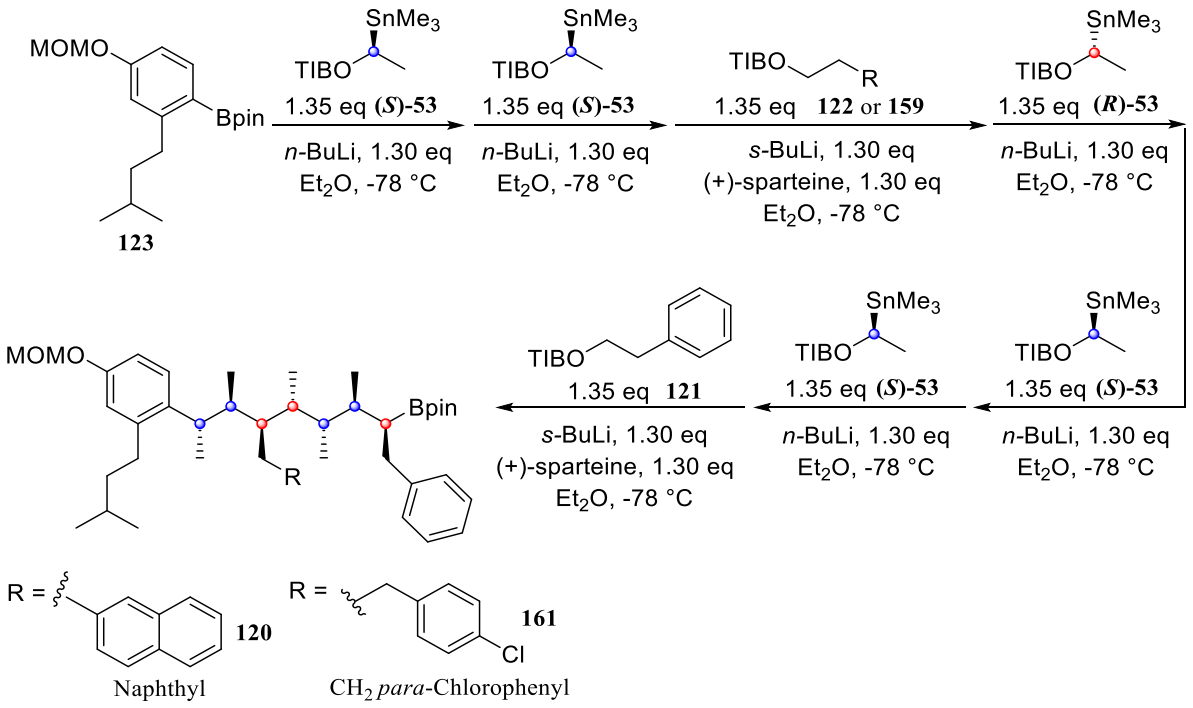


Figure 115. The results of the molecular docking study performed using AutoDock Vina. It can clearly be seen that the Cl atom of ligand **106** fills a deep void in the Trp23 pocket of Mdm2

Following the design of two suitable p53 mimetics their synthesis was explored using lithiation–borylation. In order to introduce a handle for functionalisation to improve solubility, a phenol group was added to the scaffold. The synthesis of the p53 mimetics (**157** and **164**) was performed semi-iteratively with column chromatography performed after certain homologations. The same synthetic methodology could be followed for the synthesis of both **157** and **164**, with a change to the benzoate ester used in the third homologation (**Scheme 29**). The starting boronic ester used in the synthesis of both **157** and **164** could be synthesised in four steps on a multigram scale with an overall yield of 85% obtained. From this starting boronic ester, seven homologations were performed. When installing the methyl groups, the chiral carbenoids used in the synthesis were derived from the lithiation of enantioenriched stannanes. When the hot-spot groups were installed, the chiral carbenoids were derived from the chiral-ligand assisted lithiation of primary benzoate esters. For the p53 mimetic **157**, only the first two homologations were performed iteratively, and an overall yield of 29% was obtained for the

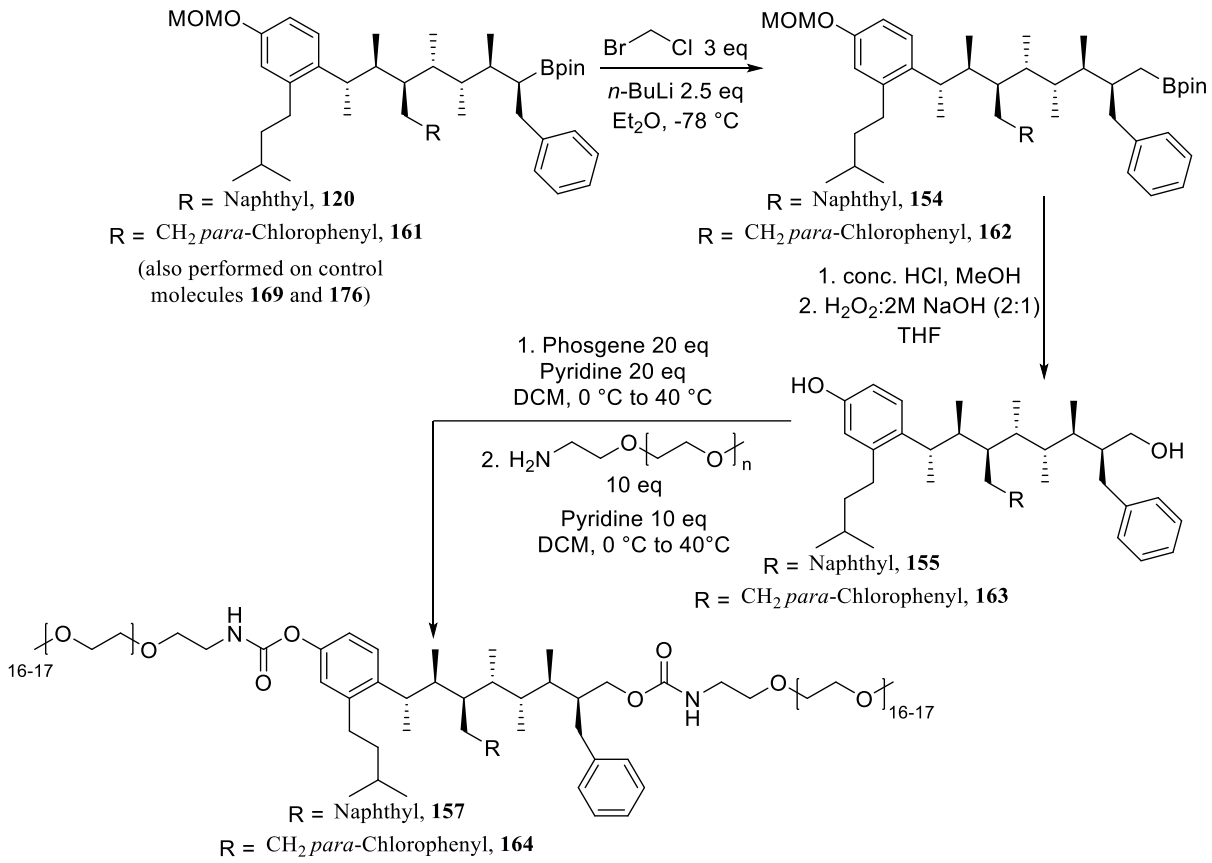
seven homologations. For the p53 mimetic **164**, the first, second, fourth, fifth and sixth homologations were performed iteratively, and an overall yield of 24% was obtained over the seven homologations.



Scheme 29. The general lithiation–borylation route followed for the synthesis of the boronic ester precursors, **120** and **161**, of the p53 mimetics **157** and **164**

In addition to the two p53 mimetics **157** and **164**, two control molecules were synthesised. Control molecule **172** was synthesised to confirm the importance of the hot-spot groups and control molecule **177** was synthesised to confirm the importance of conformational preorganisation. An iterative lithiation–borylation sequence was used to synthesise both **172** and **177**, with overall yields of 38% and 35% obtained respectively over the homologation sequence.

To improve the aqueous solubility of the scaffolds, a Matteson homologation was first performed on the terminal boronic ester to introduce a methylene unit. A MOM deprotection was performed to reveal the phenol followed by oxidation of the boronic ester to the primary alcohol. High yields were obtained over the three steps with yields ranging from 72% to 89% for the p53 mimetics and the control molecules. To introduce water solubilising groups, the two hydroxy groups on each molecule were converted to the corresponding acyl chloride with phosgene and pyridine in DCM. An *in-situ* reaction with methoxypolyethylene glycol amine introduced two PEG chains, of 750 average Mw, onto the scaffold (**Scheme 30**). The yields observed over these two steps were low, with yields of <15% obtained, due to poor conversion of the acyl chloride to the amide-PEG derivative. The bis-PEG derivatives were purified by reversed phase HPLC.



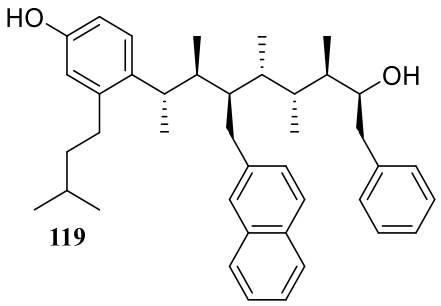
Scheme 30. The general synthesis followed for installing water solubilising PEG chains.

Of the structures synthesised, a representative was chosen on which the solution state conformation was explored (**119**). The representative structure was chosen as the candidate with the least spectral overlap to help with assignment and measurement of NMR parameters. The ^{1}H and ^{13}C chemical shifts were measured from the corresponding 1D-spectrum. Vicinal ^{1}H-^{1}H scalar couplings were measured experimentally either directly from the multiplets in the 1D-^{1}H spectrum or using SpinSimulation from MestreNova. ^{1}H-^{13}C scalar coupling constants were measured using IPAP-HSQMBC. Finally, quantitative interproton distances were measured using selective 1D-NOESY and CSSF-NOESY. The PANIC method was applied, with the irradiated peak set to 1000 and a methylene group chosen as the reference distance (1.74 Å).

The measured NMR parameters were compared to those calculated using DFT methods. A selection of conformers from the MM conformational search were subjected to a geometry optimisation and frequency calculation (mPW1PW91, 6-31g (d)), followed by a single point energy calculation (mPW1PW91, 6-311g (d,p)). The calculated conformer energies were used to calculate conformer Boltzmann populations. A final DFT calculation was performed to calculate the magnetic shielding tensors and the scalar coupling constants (GIAO, mPW1PW91, 6-311g (d,p)). The magnetic shielding tensors were converted to chemical shifts, and Boltzmann averaged using the calculated populations. The interproton distances, obtained from the optimised geometries, and the scalar coupling constants

were also Boltzmann averaged. An excellent correlation between the experimental and calculated NMR parameters was obtained, with low MADs and SDs obtained for all NMR parameters (**Table 19**). This confirms that the molecule exists in solution with a linear conformational bias.

Table 19. A summary of the MAD and SD obtained when experimental NMR parameters were compared to those calculated using DFT, for ligand **119**



119

NMR Parameter	MAD	SD
$^{n}J_{HH}$	0.71 Hz	0.58 Hz
$^{n}J_{CH}$	0.35 Hz	0.32 Hz
Interproton Distances	3.7 %	4.1 %
δ^{1}H	0.11 ppm	0.12 ppm
δ^{13}C	1.14 ppm	0.9 ppm

The final stage of this PhD was to explore the binding of the designed p53 mimetics to Mdm2. This was performed by ^{1}H-^{15}N HSQC spectroscopy. This allows the location of binding to Mdm2 to be identified as well as the dissociation constant (K_d) to be measured. ^{15}N-Mdm2 was expressed in the *E.Coli* BL21(DE3) strain using M9 minimal media with $^{15}NH_4Cl$ as the nitrogen source. Initial purification using IMAC followed by a subsequent purification using SEC yielded good quantities of ^{15}N-Mdm2. The mass was confirmed by MALDI spectroscopy, which confirmed that full ^{15}N labelling had occurred. A ^{1}H-^{15}N TROSY was acquired and compared to those reported in the literature which confirmed that Mdm2 had folded into the correct tertiary structure. A partial assignment (76%) of the peaks was achieved by comparison to the assignments in the BMRB (2410 and 15945). Upon binding of the naphthyl derivative **157**, significant CSPs were observed for residues located primarily in, or on the periphery of, the p53 binding pocket (**Figure 116**). CSPs that were observed for residues not in the p53 binding pocket were attributed to conformational changes in Mdm2 upon accommodation of a ligand. Similar CSPs were observed upon binding of the *para*-chloro phenyl derivative **164**, although the magnitude of the CSPs is greater, suggesting that **164** is binding to Mdm2 with a greater affinity.

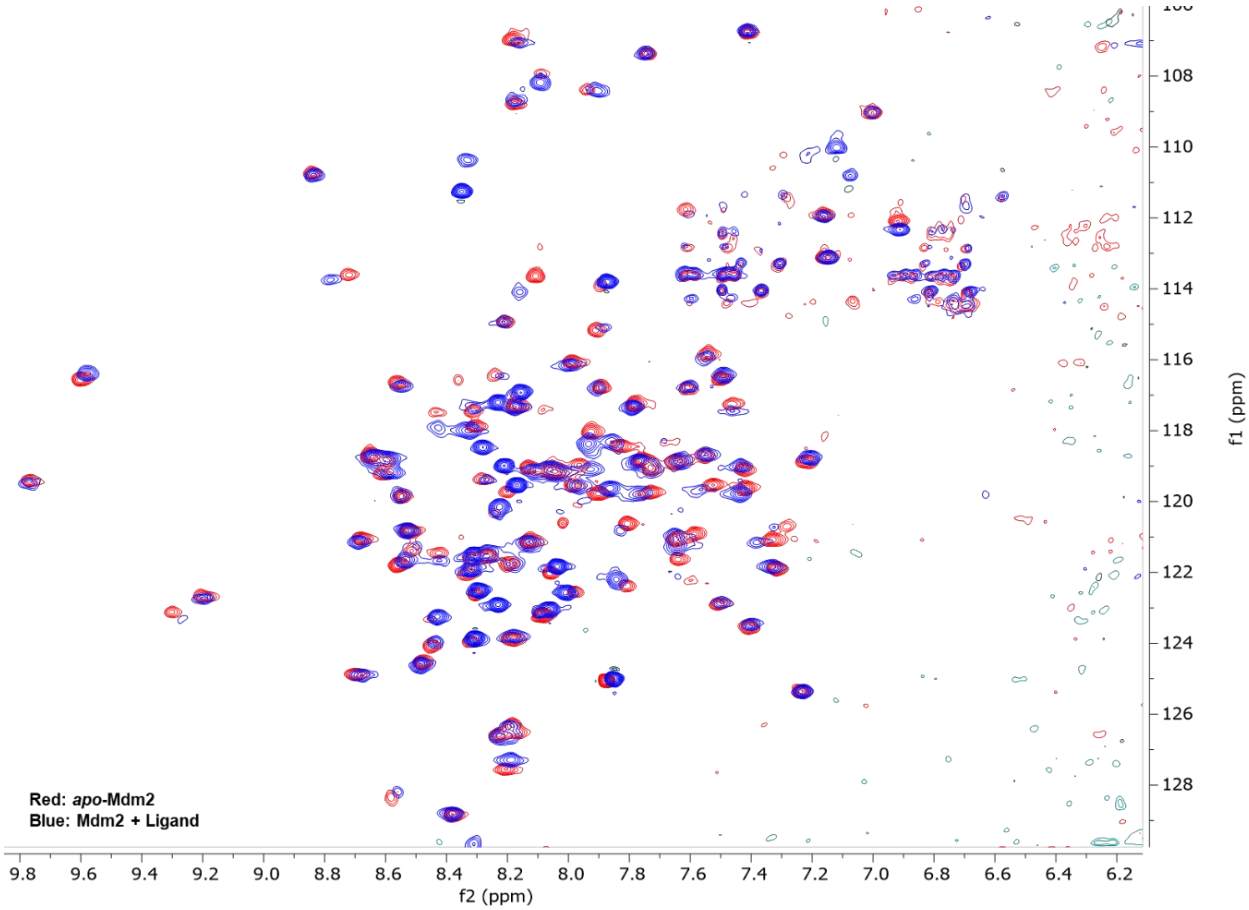
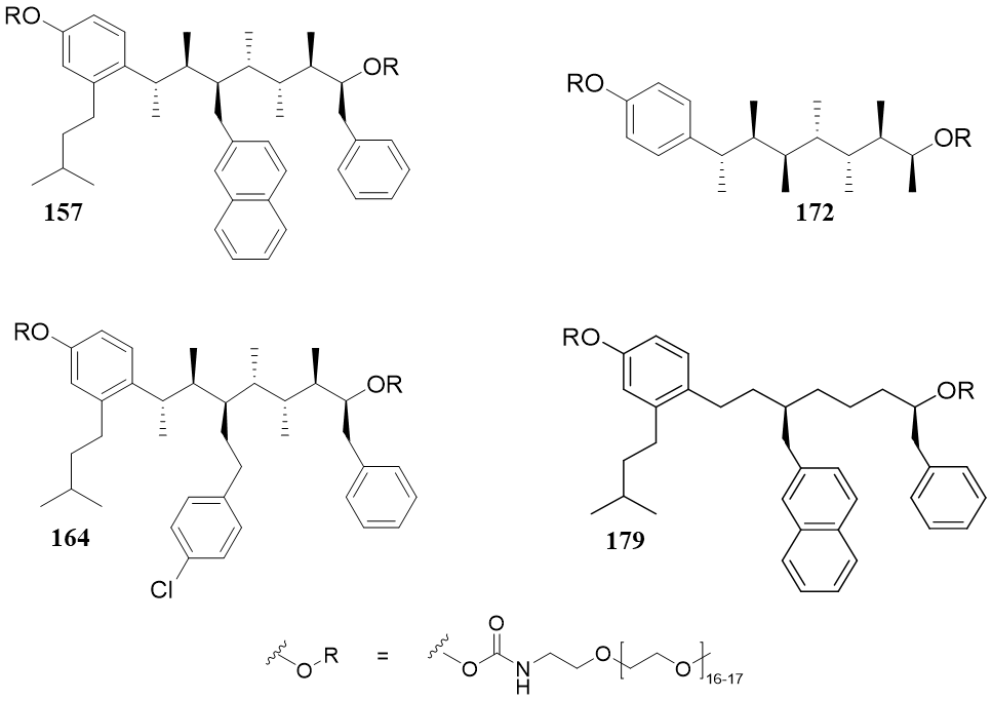


Figure 116. The ^{1}H-^{15}N TROSY obtained for *apo*-Mdm2 (red) and Mdm2+ligand **157**. CSPs can be clearly observed for certain residues which can be mapped to residues located in the p53 binding pocket of Mdm2.

To estimate the dissociation constant K_d, a titration of five increasing concentrations of the ligands **157** and **164** against Mdm2 was performed. Mbinding from MestreNova was used to analyse the ^{1}H-^{15}N SOFAST-HMQC spectra. Mbinding calculates the Euclidean CSP at each concentration, plots the calculated binding curves and estimates the K_d. In addition to the two p53 mimetics **157** and **164**, the binding of the two control molecules **172** and **179** was explored. A summary of the K_d values obtained for the two p53 mimetics **157** and **164** from ^{1}H-^{15}N SOFAST-HMQC spectroscopy is shown in **Table 20**. It confirms that the *para*-chloro phenyl derivative (**164**) binds to Mdm2 with a higher affinity than the naphthyl derivative (**157**), as predicted by molecular docking. The two control molecules, **172** and **179**, do not induce significant CSPs and reliable values of K_d cannot be obtained. However, the lack of CSPs confirms that the two control molecules bind to Mdm2 with a considerably weaker affinity than ligands **157** and **164**, confirming the importance of both the hot-spot groups and conformational preorganisation for obtaining higher affinity therapeutics in the inhibition of PPIs.

Table 20. A summary of the dissociation constants (K_d) obtained from ^{1}H-^{15}N HSQC titrations for the two p53 mimetics



Ligand	K_d (μM)
157	17.18
164	8.81

In summary, a new class of conformationally controlled α-helix mimetics has been designed and synthesised. Using our modular and flexible iterative lithiation–borylation methodology, a vast range of groups can be incorporated onto the scaffold through the simple substitution of the benzoate ester used in the synthetic sequence. This thesis has shown that these scaffolds can mimic one face of the α-helix (the *i*, *i* + 3/4 and *i* + 7) however, it is also believed that they can be used as multi-faced α-helix mimetics and this will be explored in due course. Finally, it was demonstrated that the scaffold can be functionalised with suitable groups to mimic the Mdm2 binding domain to p53 and bind to the p53 binding pocket on Mdm2 with binding affinities as low as 9 μM. We hope that this work will encourage the exploration of this new conformationally controlled α-helix mimetic in the inhibition of further aberrant PPIs.

Chapter 9: Supporting Information

9.1 Molecular Mechanics Conformational Searching

9.1.1 General Informational

Molecular Mechanics Conformational Search Calculations were performed using the MacroModel[322] Software package (version 9.9) accessed through Maestro (version 9.2)[323]. All MM calculations were performed using Grendel, the Unix computational resource of the School of Chemistry, University of Bristol.

Conformational searching was performed using the Monte Carlo Multiple Minimum (MCMM)[120] conformational search method. The Merck Molecular Force Field (MMFFs)[236] was used with 500,000 steps in the gas phase. Conformers within 21 kJ mol^{-1} of the global energy minima were retained. Structures were minimised using the Truncated Newton Conjugate Gradient (TNCG) method with 1000 iterations and a convergence criterion of 0.05 Å. If poor convergence was observed, noted by a low number of times the low energy conformers were found, the number of iterations was increased to 5000.

Following the conformational search, a redundant conformer elimination was performed using MacroModel. In this study, conformers are deemed redundant and eliminated using maximum atom deviations. This considers a conformer to be unique if the maximum atom deviation for any pair of corresponding atoms exceeds 0.5 Å. The atoms along the backbone and the first carbon atom of the side chains were selected for comparison (**Figure 117**). All conformers under 21 kJ mol^{-1} were retained.

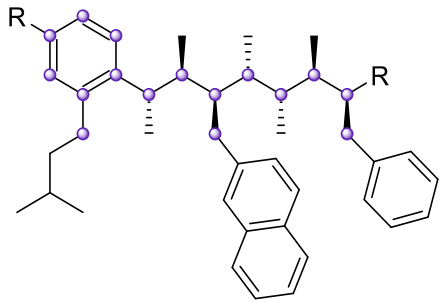


Figure 117. The atoms considered for comparison during the redundant conformer elimination.

To generate the bubble plots, the dihedral angles along the backbone of each conformer were measured in MacroModel. The estimated relative potential energies of each conformer from MM was used to calculate the Boltzmann distribution of conformers according to **Equation 13**. A 3D plot was then created using the measured dihedral angles, the dihedral number and the conformer populations.

$$p_i = \frac{g e^{-\frac{\Delta E}{RT}}}{\sum_{i=1}^{n} g e^{-\frac{\Delta E}{RT}}}$$

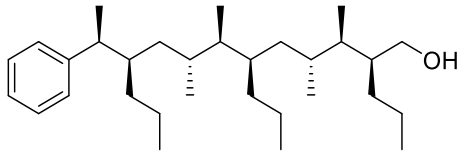
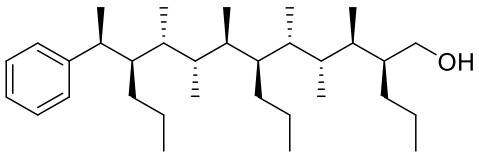
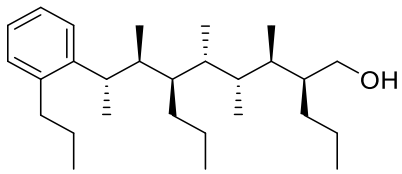
Equation 13. The Boltzmann Distribution

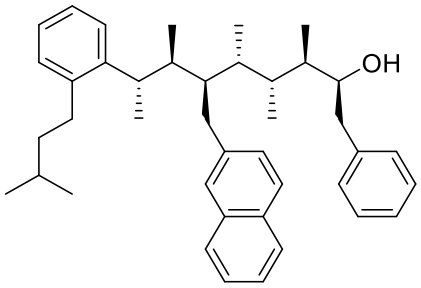
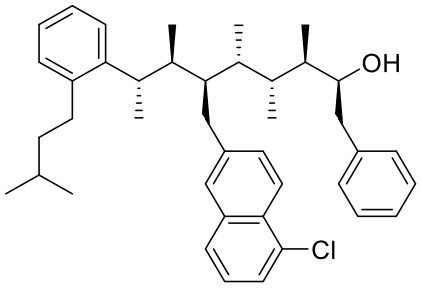
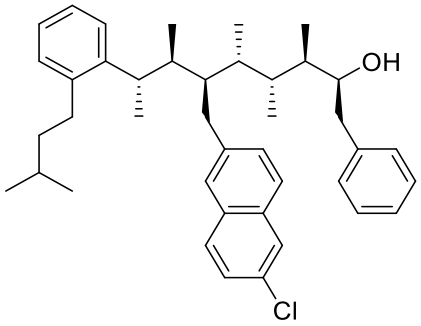
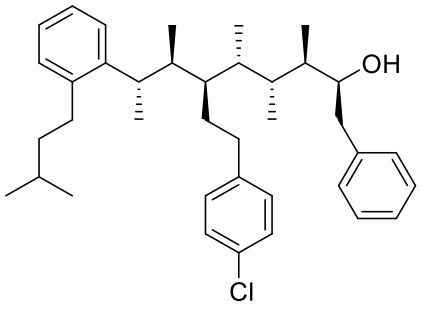
Where ΔE is the estimated potential energy, relative to the global minimum, in kJ mol^{-1} of conformer *i*. R is the ideal gas constant (8.314 x 10 $^{-3}$ kJ mol^{-1}), T is the temperature of the system (298 K) and *g* is the degeneracy of conformer *i* (which is 1 for all conformers studied).

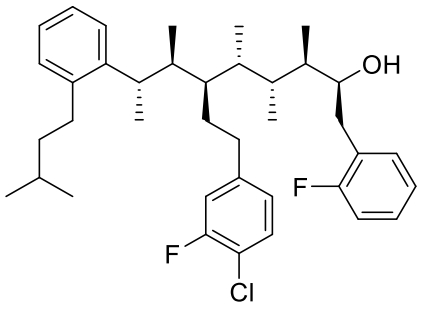
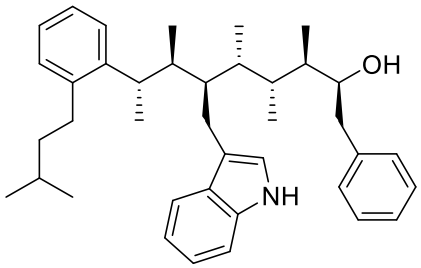
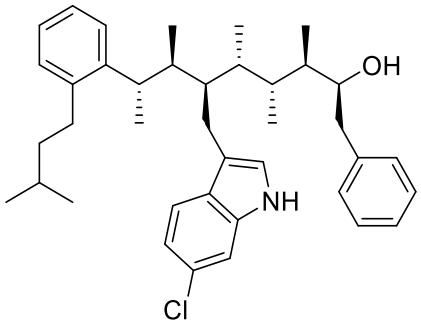
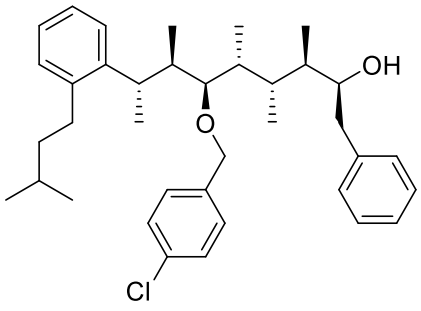
9.1.2 Conformational Search Results of Initial Candidates

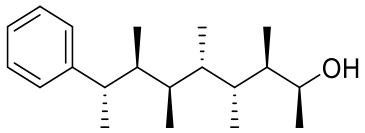
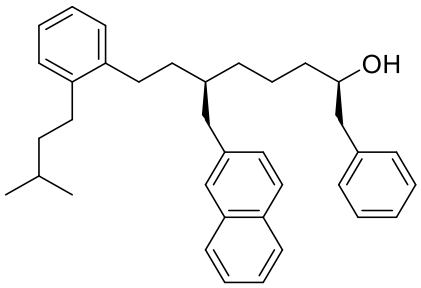
A summary of the conformational search results of the structures investigated during the computational design stage of this report are shown in **Table 21**.

Table 21. A summary of the conformational search results for all structures considered during the computational design of a p53 mimetic.

Structure	Number of Conformers within 21 kJ mol^{-1}	Number of Conformers after Redundant Conformer Elimination	% of Linear Structures.
 **100**	2914	276	0%
 **101**	2716	198	73%
 **102**	2088	74	94%

Structure	Number of Conformers within 21 kJ mol^{-1}	Number of Conformers after Redundant Conformer Elimination	% of Linear Structures.
 **96**	3147	118	74%
 **103**	4172	292	75%
 **104**	4005	281	69%
 **105**	6656	369	80%

Structure	Number of Conformers within 21 kJ mol^{-1}	Number of Conformers after Redundant Conformer Elimination	% of Linear Structures.
 **106**	7057	399	72%
 **107**	2586	244	65%
 **108**	2764	160	58%
 **110**	2868	156	67%

Structure	Number of Conformers within 21 kJ mol^{-1}	Number of Conformers after Redundant Conformer Elimination	% of Linear Structures.
 **111**	43	n/a	97%
 **112**	44904	3049	1.4%

9.1.3 Conformational Search Results of Synthesised Structures

A summary of the conformational search results of 119

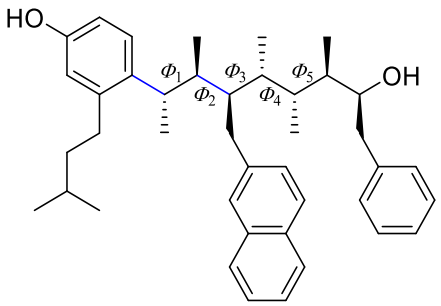


Table 22. The MM conformational search results of **119**.

Number of Conformers within 30 kJ mol^{-1}	Number of Conformers after Redundant Conformer Elimination	% of Linear Structures.
17380	1115	77%

The lowest energy conformer of **119** is linear and has a population of 22%.

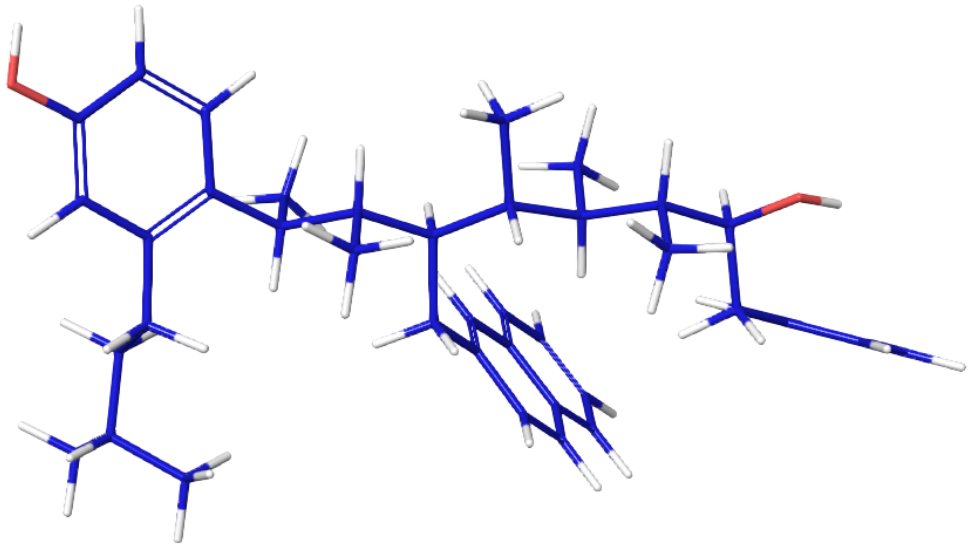


Figure 118. The lowest energy conformer of **119** from MM.

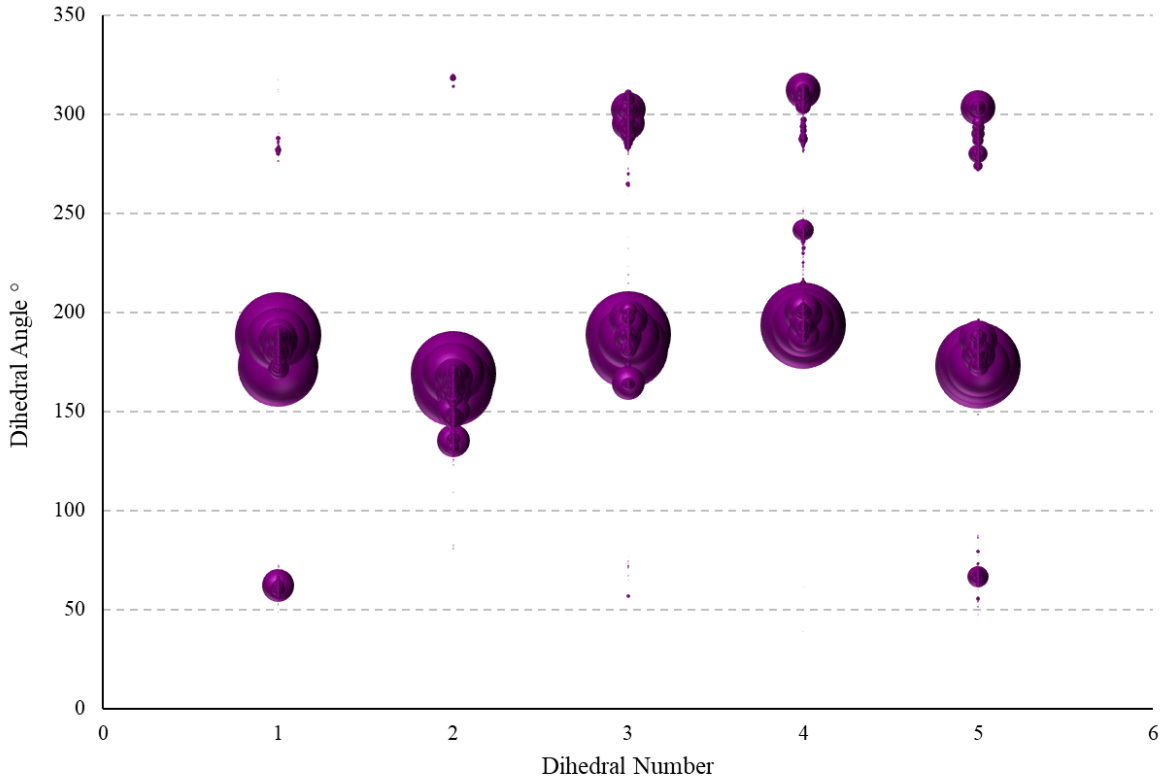


Figure 119. The Bubble plot of **119.** The size of the bubble is representative of the population of the conformer contributing to that dihedral angle.

A summary of the conformational search results of 155

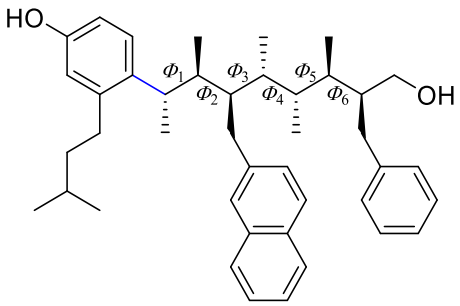


Table 23. The MM conformational search results of **155**

Number of Conformers within 21 kJ mol^{-1}	Number of Conformers after Redundant Conformer Elimination	% of Linear Structures.
9577	200	72%

The lowest energy conformer of **155** is linear and has a population of 21%.

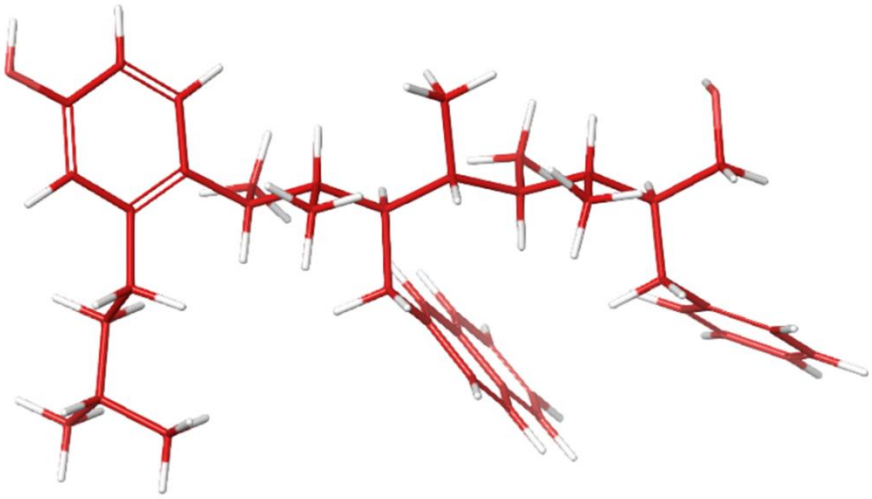


Figure 120. The lowest energy conformer from MM for **155**

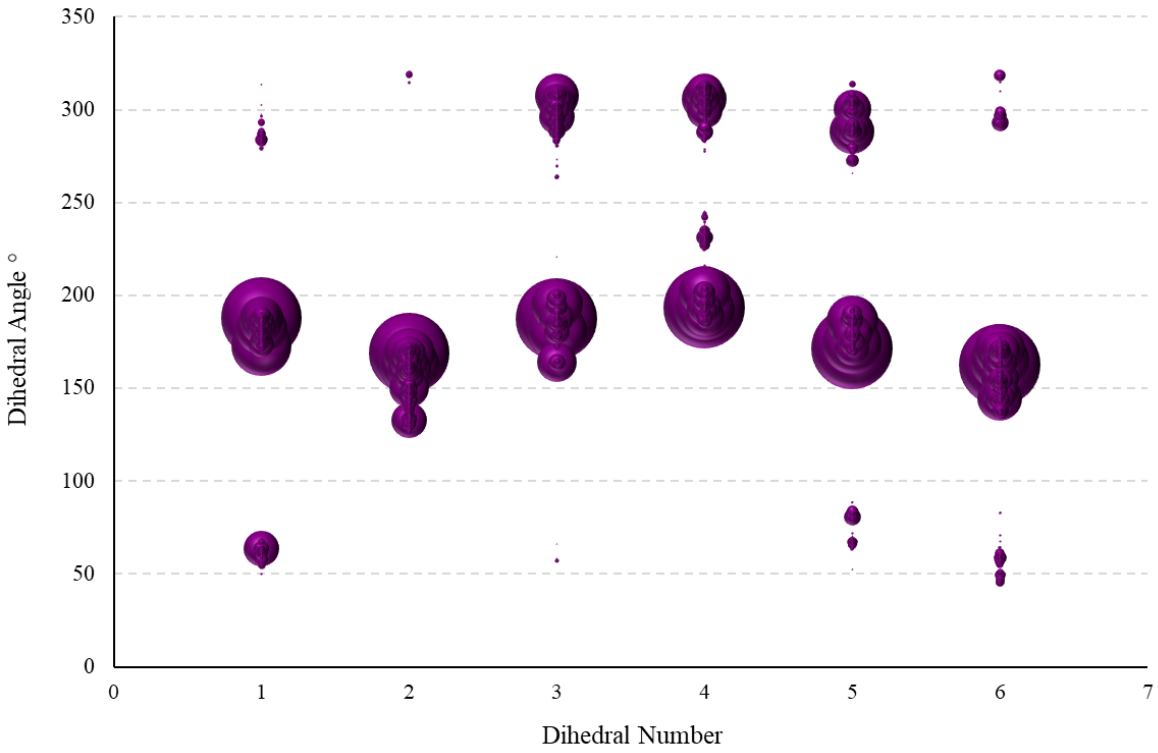


Figure 121. The Bubble plot of **155**. The size of the bubble is representative of the population of the conformer contributing to that dihedral angle.

A summary of the conformational search results of 163

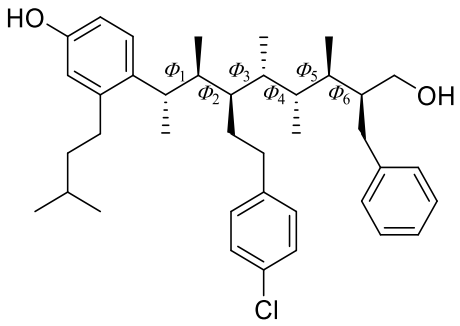


Table 24. The MM conformational search results of **163**

Number of Conformers within 21 kJ mol^{-1}	Number of Conformers after Redundant Conformer Elimination	% of Linear Structures.
6277	188	75%

The lowest energy conformer of **163** is linear and has a population of 37%.

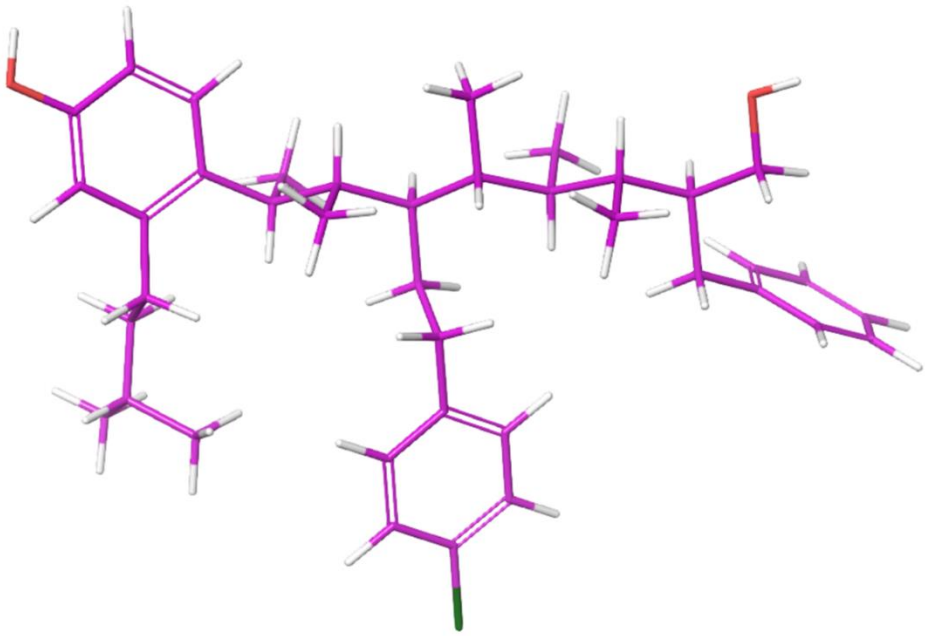


Figure 122. The lowest energy conformer from MM for **163**

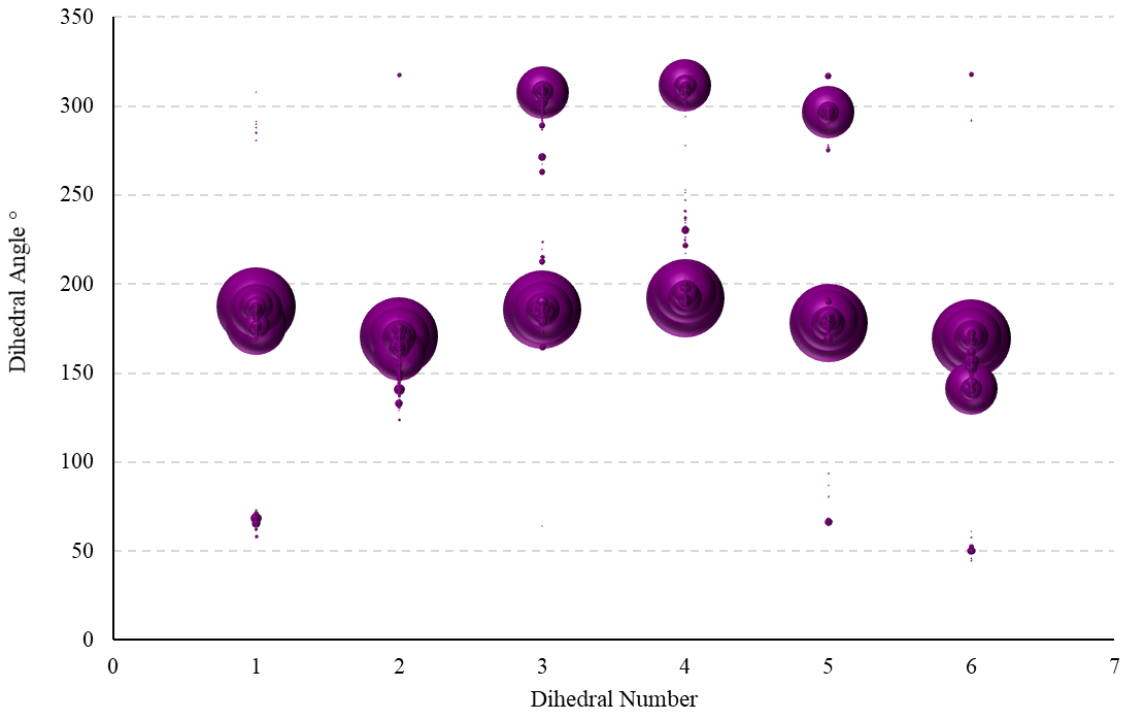


Figure 123. The Bubble plot of **163**. The size of the bubble is representative of the population of the conformer contributing to that dihedral angle.

A summary of the conformational search results of 171

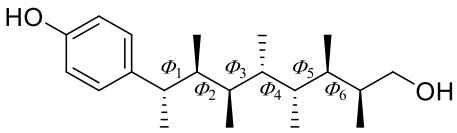


Table 25. The MM conformational search results of **171**

Number of Conformers within 21 kJ mol^{-1}	Number of Conformers after Redundant Conformer Elimination	% of Linear Structures.
130	n/a	98%

The lowest energy conformer of **171** is linear and has a population of 13%

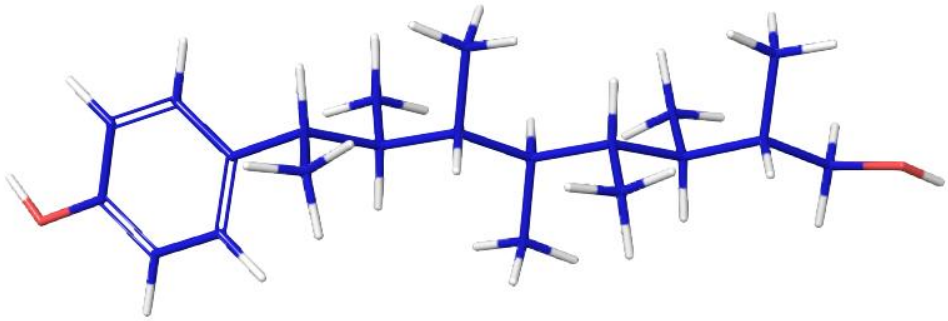


Figure 124. The lowest energy conformer from MM for **171**

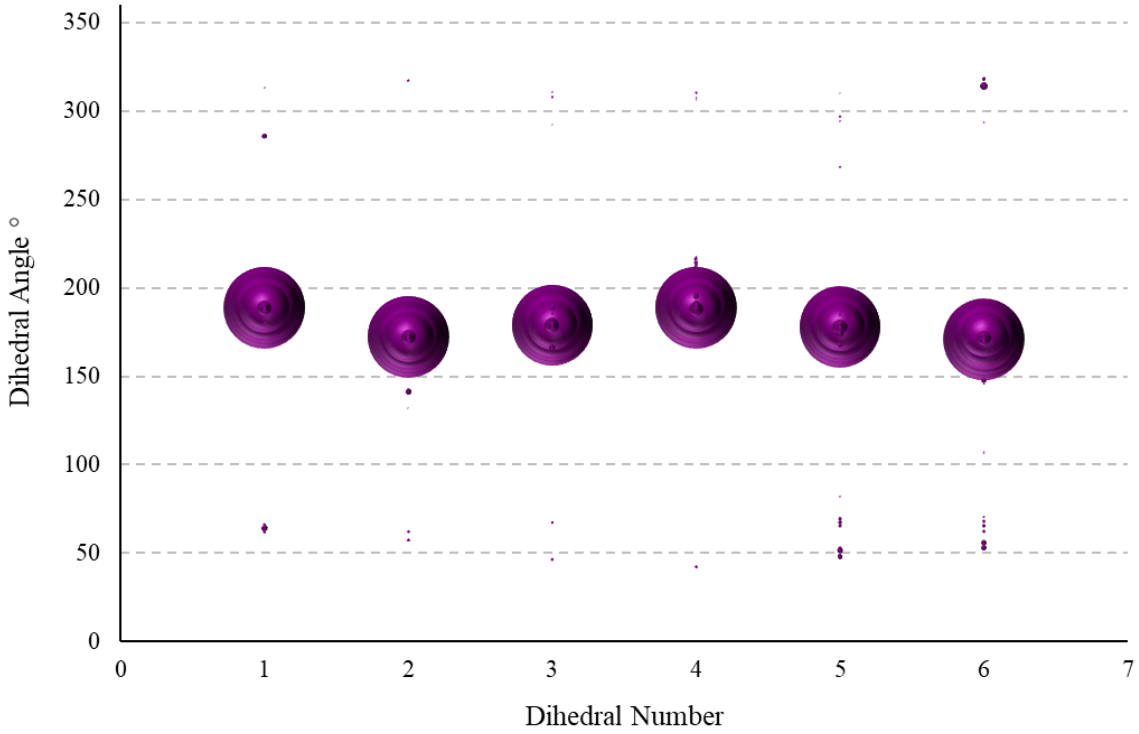


Figure 125. The Bubble plot of **171**. The size of the bubble is representative of the population of the conformer contributing to that dihedral angle.

A summary of the conformational search results of 178

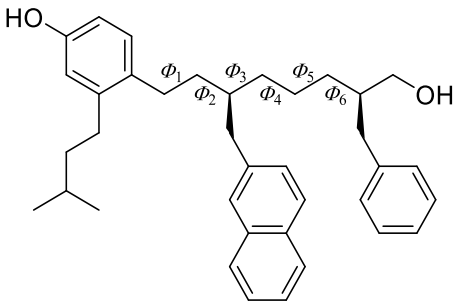


Table 26. The MM conformational search results of **178**

Number of Conformers within 21 kJ mol^{-1}	Number of Conformers within 10 kJ mol^{-1} after Redundant Conformer Elimination*	% of Linear Structures.
86446	1581	<0.1%

*only atoms of the backbone were selected. Side chains were ignored

The lowest energy conformer of **178** is not linear and has a population of 5%

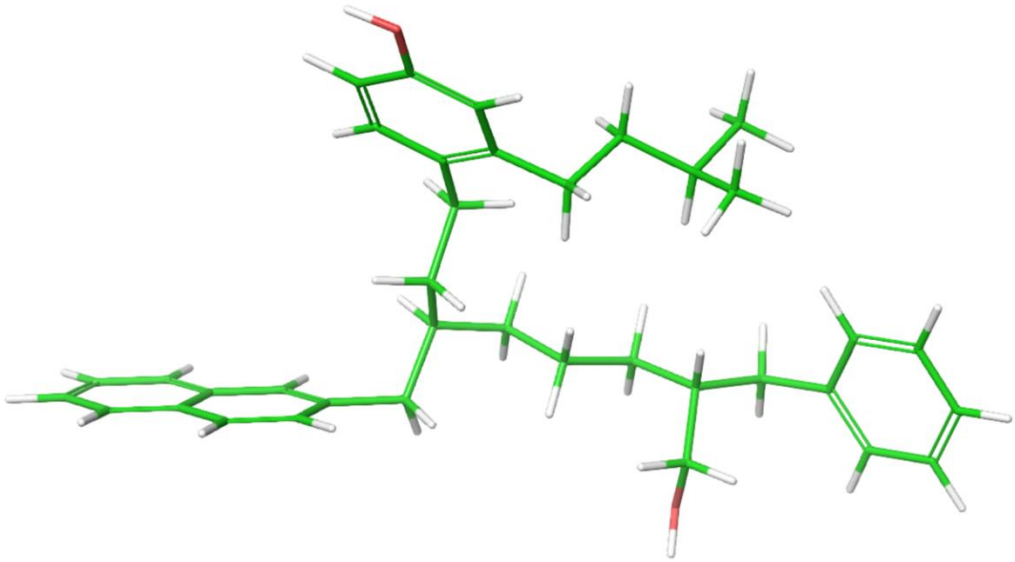


Figure 126. The lowest energy conformer from MM for **178**

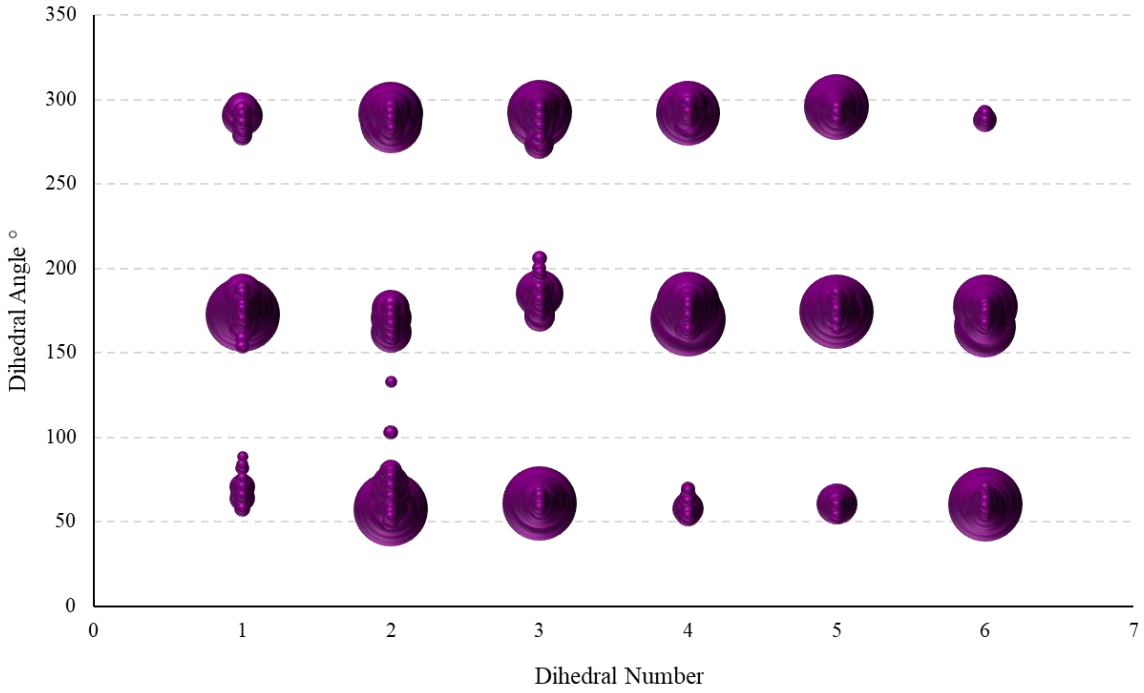


Figure 127. The Bubble plot of **178.** The size of the bubble is representative of the population of the conformer contributing to that dihedral angle.

9.2 Molecular Docking using AutoDock Vina

Molecular Docking was performed using AutoDock Vina, version 1.1.2.[243] All procedure for the docking experiment were followed as described in the user manual for AutoDock Vina with the exception of the exhaustiveness, which was increased to 24. The ligand and receptor were parameterised using Autodock Tools. Docked conformations were ranked automatically by AutoDock Vina using a force field scoring function. The parameters (centre point of binding and size of the search space) used in the calculation are:

center_x = 30.804
center_y = -18.102
center_z = -3.946
size_x = 30.0
size_y = 30.0
size_z = 24.0

The binding of Nutlin-2 by AutoDock Vina

Mdm2 was taken from the crystal structure 1YCR and the lowest energy conformer of Nutlin-2, found from a MM conformational search, was used as the ligand. The best binding mode gave a predicted binding affinity of -41.4 kJ mol^{-1}.

Table 27. The top five binding poses of Nutlin-2 bound to Mdm2 found by AutoDock-Vina.

Mode	Affinity (kJ mol^{-1})	Distance from Best Mode (Å)
1	-41.42	0.00
2	-39.75	7.94
3	-37.66	7.46
4	-33.89	4.95
5	-33.05	7.15

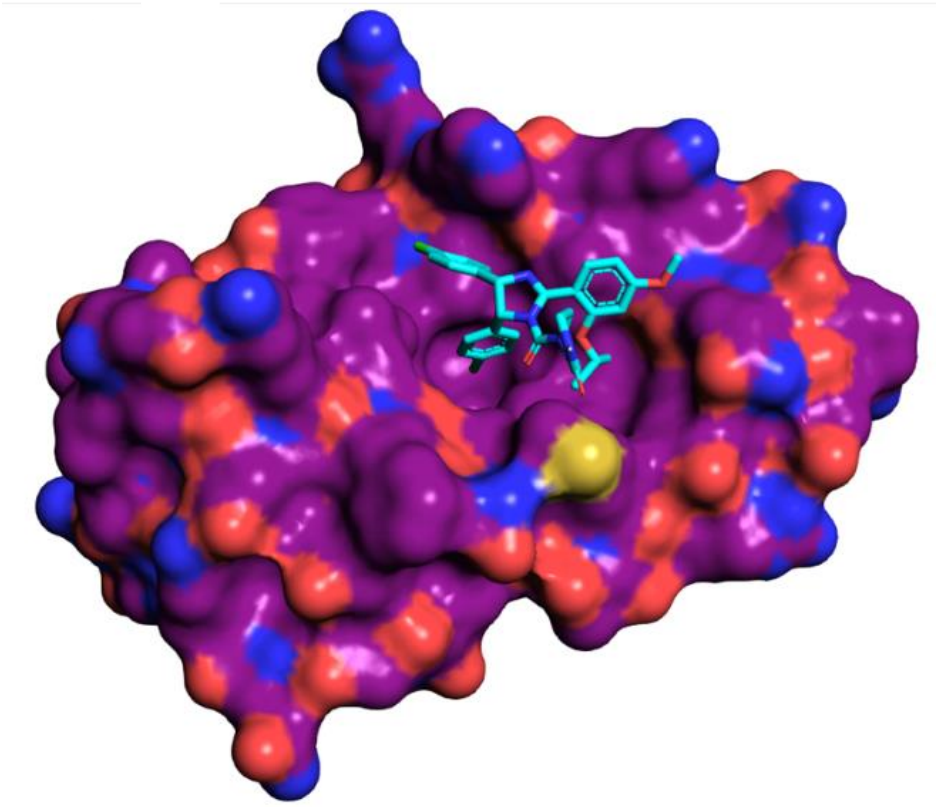
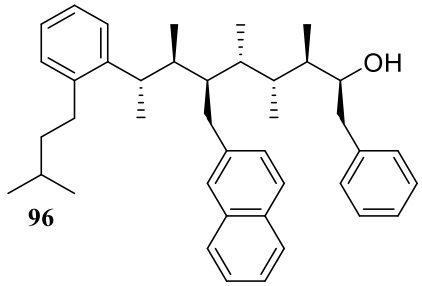


Figure 128. The top ranked binding pose of Nutlin-2 bound to Mdm2 found by AutoDock Vina.

The binding of 96 by AutoDock Vina



96

Mdm2 was taken from the crystal structure 1YCR and the lowest energy conformer of **96**, found from a MM conformational search, was used as the ligand. The best binding mode gave a predicted binding affinity of -36.8 kJ mol^{-1}.

Table 28. The top five binding poses of **96** bound to Mdm2 found by AutoDock-Vina.

Mode	Affinity (kJ mol^{-1})	Distance from Best Mode (Å)
1	-36.82	0.00
2	-34.31	2.45
3	-32.64	11.19
4	-32.22	8.13
5	-31.80	9.69

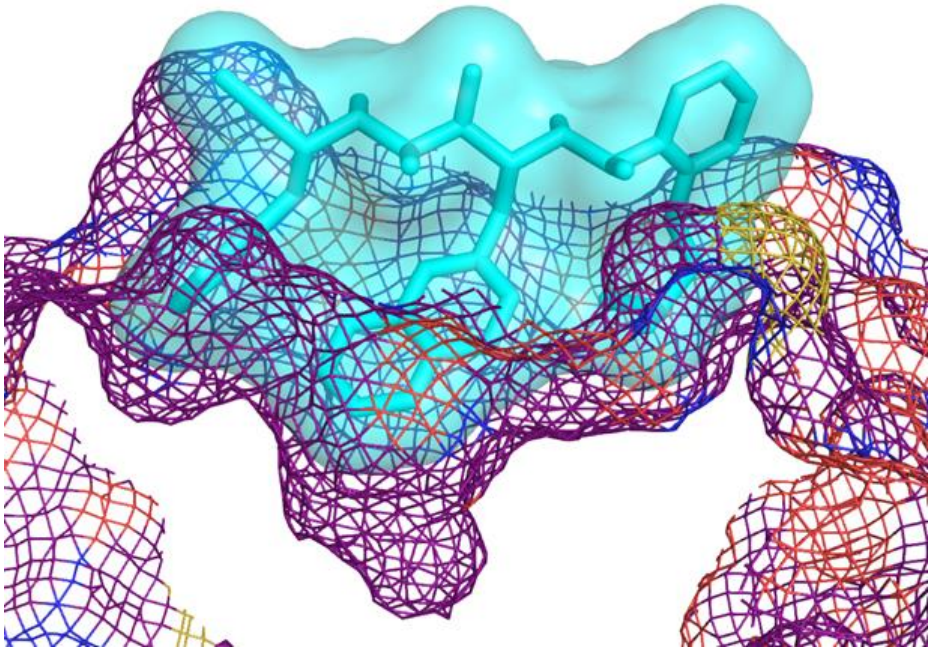
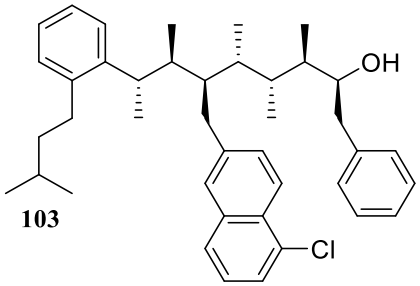


Figure 129. The top ranked binding pose of **96** bound to Mdm2 found by AutoDock Vina. A surface representation of **96** bound to Mdm2 is shown to display the occupancy of the p53 pocket on Mdm2 by **96**.

The binding of 103 by AutoDock Vina



103

Mdm2 was taken from the crystal structure 1YCR and the lowest energy conformer of **103**, found from a MM conformational search, was used as the ligand. The best binding mode gave a predicted binding affinity of -34.73 kJ mol^{-1}.

Table 29. The top five binding poses of **103** bound to Mdm2 found by AutoDock-Vina.

Mode	Affinity (kJ mol^{-1})	Distance from Best Mode (Å)
1	-34.73	0.00
2	-32.64	3.98
3	-32.22	2.49
4	-31.38	4.50
5	-30.96	3.30

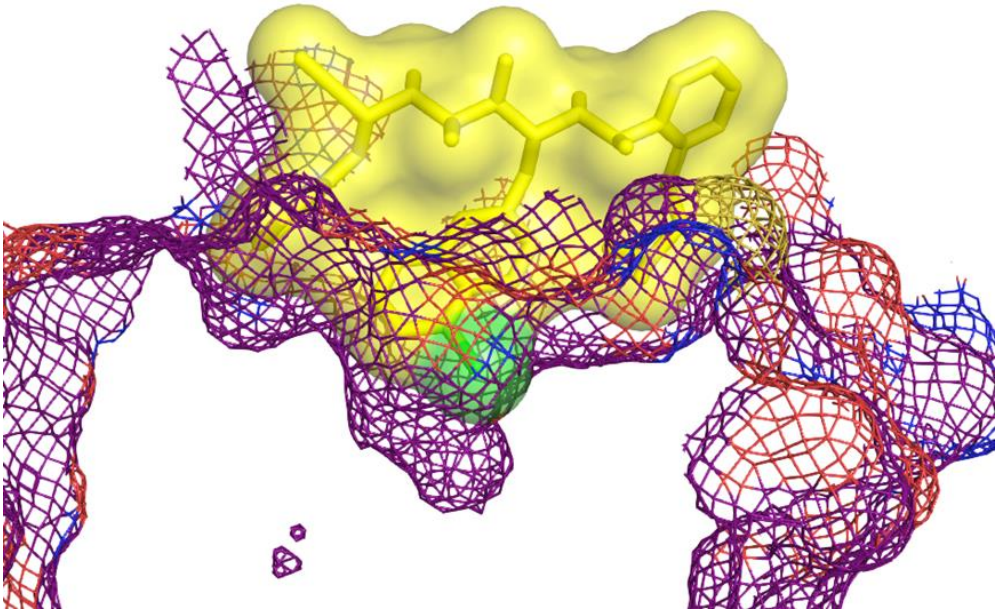
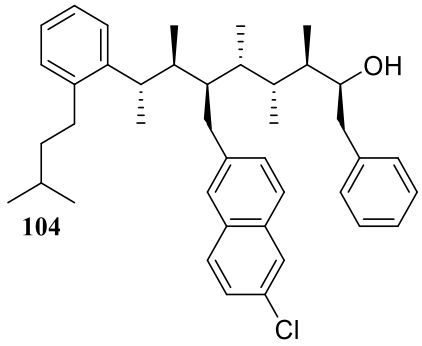


Figure 130. The top ranked binding pose of **103** bound to Mdm2 found by AutoDock Vina. A surface representation of **103** bound to Mdm2 is shown to display the occupancy of the p53 pocket on Mdm2 by **103**.

The binding of 104 by AutoDock Vina



104

Mdm2 was taken from the crystal structure 1YCR and the lowest energy conformer of **104**, found from a MM conformational search, was used as the ligand. The best binding mode gave a predicted binding affinity of -37.24 kJ mol^{-1}.

Table 30. The top five binding poses of **104** bound to Mdm2 found by AutoDock-Vina.

Mode	Affinity (kJ mol^{-1})	Distance from Best Mode (Å)
1	-37.24	0.00
2	-36.40	2.60
3	-35.15	3.89
4	-34.73	2.58
5	-34.73	2.39

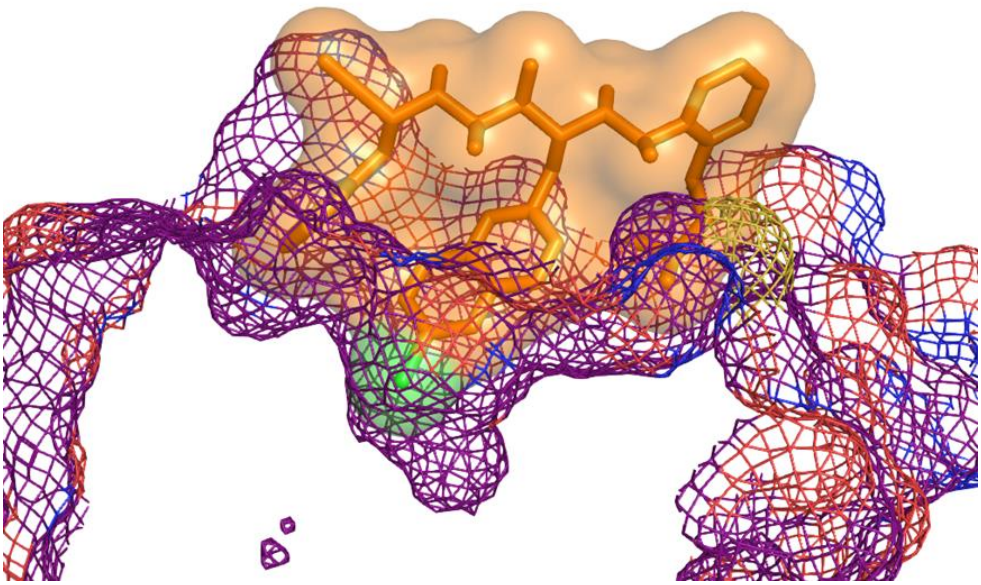
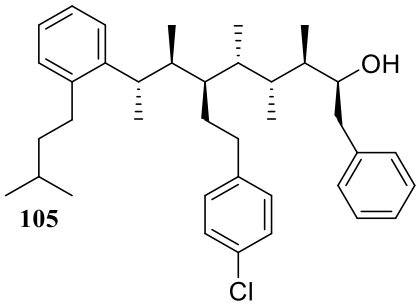


Figure 131. The top ranked binding pose of **104** bound to Mdm2 found by AutoDock Vina. A surface representation of **104** bound to Mdm2 is shown to display the occupancy of the p53 pocket on Mdm2 by **104**.

The binding of 105 by AutoDock Vina



Mdm2 was taken from the crystal structure 1YCR and the lowest energy conformer of **105**, found from a MM conformational search, was used as the ligand. The best binding mode gave a predicted binding affinity of -38.9 kJ mol^{-1}.

Table 31. The top five binding poses of **105** bound to Mdm2 found by AutoDock-Vina.

Mode	Affinity (kJ mol^{-1})	Distance from Best Mode (Å)
1	-38.91	0.00
2	-37.24	2.38
3	-35.98	2.23
4	-35.56	2.59
5	-33.89	4.33

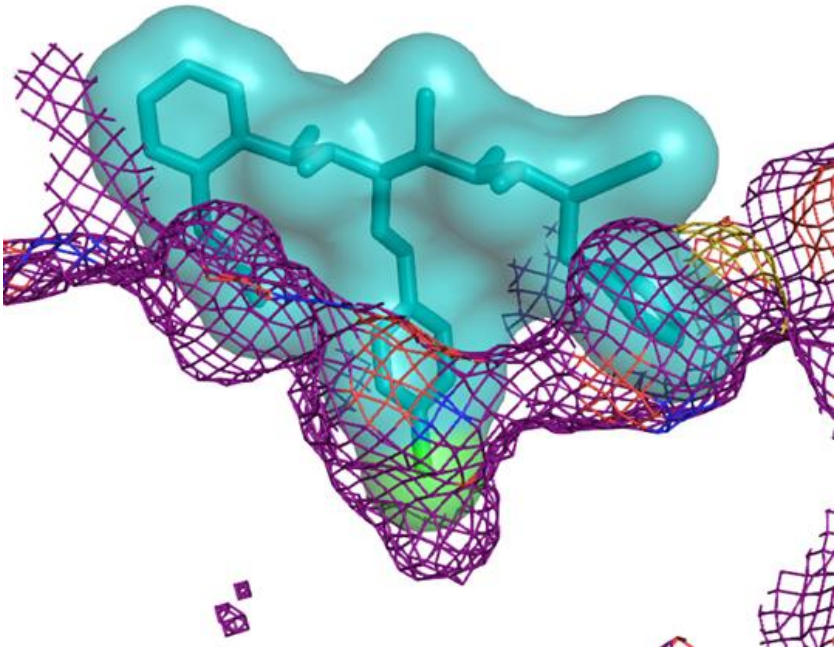
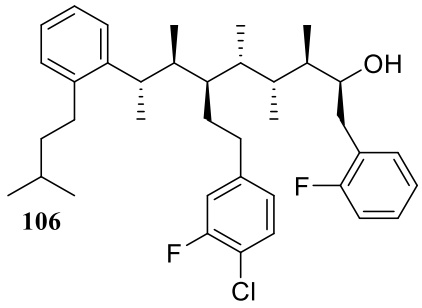


Figure 132. The top ranked binding pose of **105** bound to Mdm2 found by AutoDock Vina. A surface representation of **105** bound to Mdm2 is shown to display the occupancy of the p53 pocket on Mdm2 by **105**.

The binding of 106 by AutoDock Vina



106

Mdm2 was taken from the crystal structure 1YCR and the lowest energy conformer of **106**, found from a MM conformational search, was used as the ligand. The best binding mode gave a predicted binding affinity of -41.00 kJ mol^{-1}.

Table 32. The top five binding poses of **106** bound to Mdm2 found by AutoDock-Vina.

Mode	Affinity (kJ mol^{-1})	Distance from Best Mode (Å)
1	-41.00	0.00
2	-38.07	2.21
3	-36.82	3.36
4	-36.40	1.14
5	-35.56	2.58

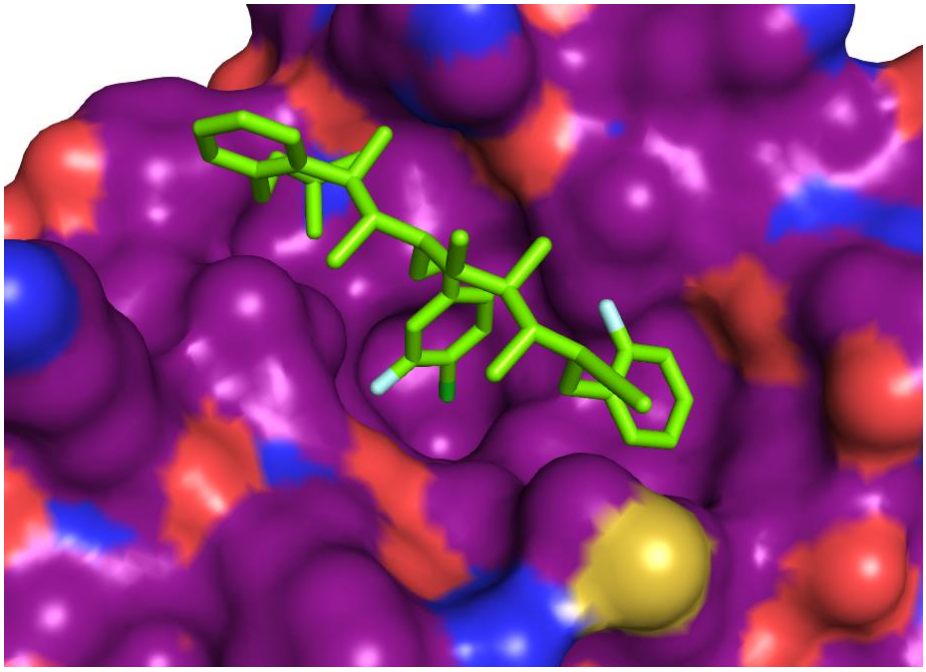
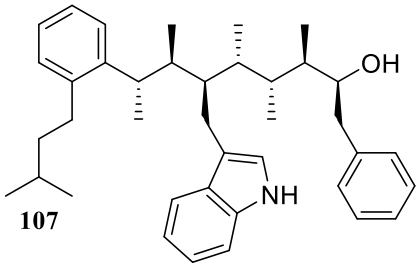


Figure 133. The top ranked binding pose of **106** bound to Mdm2 found by AutoDock Vina.

The binding of 107 by AutoDock Vina



Mdm2 was taken from the crystal structure 1YCR and the lowest energy conformer of **107**, found from a MM conformational search, was used as the ligand. The best binding mode gave a predicted binding affinity of -35.15 kJ mol^{-1}.

Table 33. The top five binding poses of **107** bound to Mdm2 found by AutoDock-Vina.

Mode	Affinity (kJ mol^{-1})	Distance from Best Mode (Å)
1	-35.15	0.00
2	-35.15	2.21
3	-32.64	3.57
4	-32.22	4.40
5	-32.22	2.78

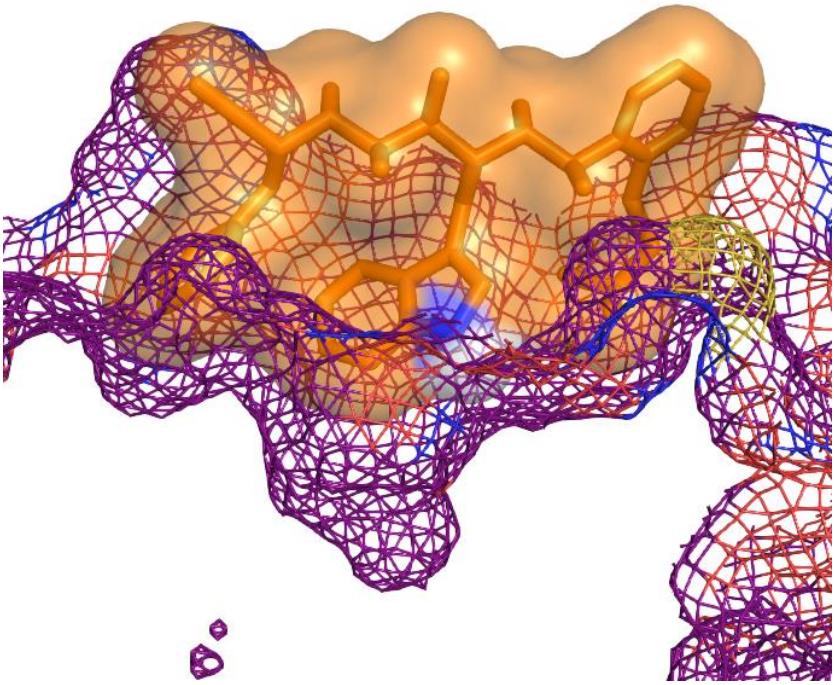
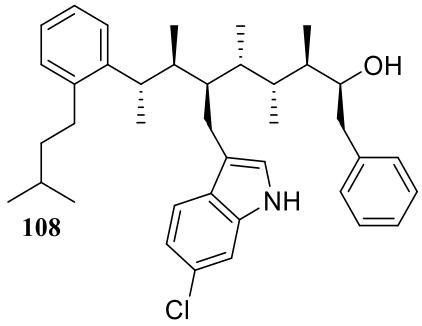


Figure 134. The top ranked binding pose of **107** bound to Mdm2 found by AutoDock Vina. A surface representation of **107** bound to Mdm2 is shown to display the occupancy of the p53 pocket on Mdm2 by **107**.

The binding of 108 by AutoDock Vina



108

Mdm2 was taken from the crystal structure 1YCR and the lowest energy conformer of **108**, found from a MM conformational search, was used as the ligand. The best binding mode gave a predicted binding affinity of -37.7 kJ mol^{-1}.

Table 34. The top five binding poses of **108** bound to Mdm2 found by AutoDock-Vina.

Mode	Affinity (kJ mol^{-1})	Distance from Best Mode (Å)
1	-37.66	0.00
2	-36.82	1.76
3	-36.40	1.91
4	-36.40	4.88
5	-35.98	2.43

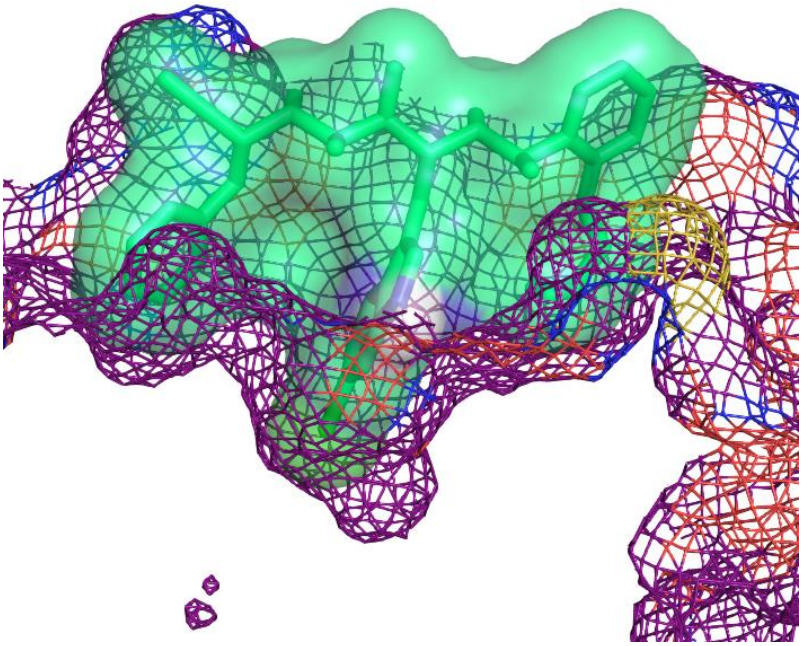
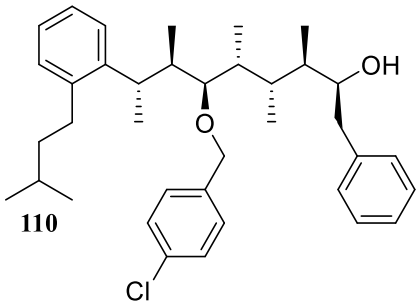


Figure 135. The top ranked binding pose of **108** bound to Mdm2 found by AutoDock Vina. A surface representation of **108** bound to Mdm2 is shown to display the occupancy of the p53 pocket on Mdm2 by **108**.

The binding of 110 by AutoDock Vina



110

Mdm2 was taken from the crystal structure 1YCR and the lowest energy conformer of **110**, found from a MM conformational search, was used as the ligand. The best binding mode gave a predicted binding affinity of -38.07 kJ mol^{-1}.

Table 35. The top five binding poses of **110** bound to Mdm2 found by AutoDock-Vina.

Mode	Affinity (kJ mol^{-1})	Distance from Best Mode (Å)
1	-38.07	0.00
2	-35.15	2.44
3	-35.15	2.29
4	-34.31	2.19
5	-33.89	4.63

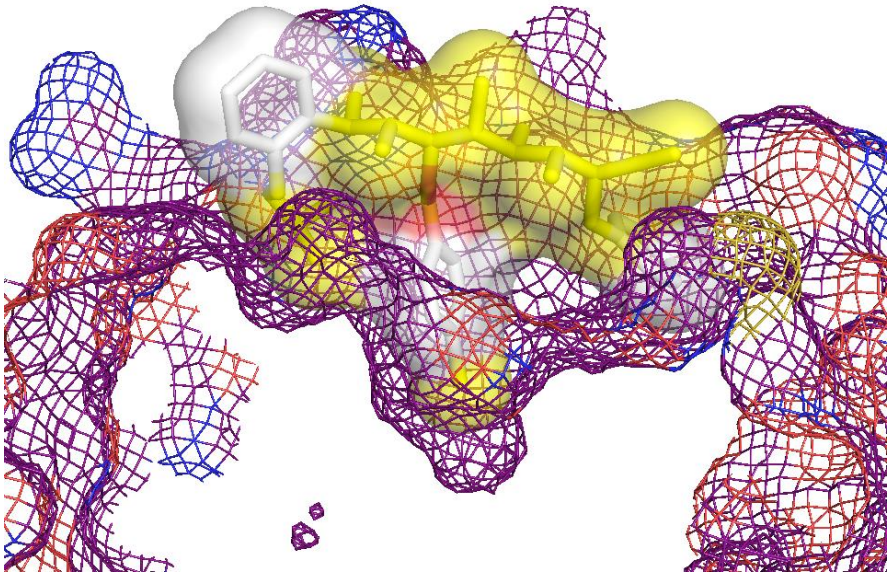
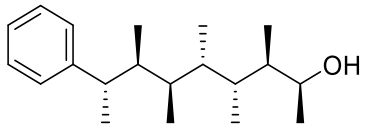


Figure 136. The top ranked binding pose of **110** bound to Mdm2 found by AutoDock Vina. A surface representation of **110** bound to Mdm2 is shown to display the occupancy of the p53 pocket on Mdm2 by **110**.

The binding of 111 by AutoDock Vina



111

Mdm2 was taken from the crystal structure 1YCR and the lowest energy conformer of **111**, found from a MM conformational search, was used as the ligand. The best binding mode gave a predicted binding affinity of -29.3 kJ mol^{-1}.

Table 36. The top five binding poses of **111** bound to Mdm2 found by AutoDock-Vina

Mode	Affinity (kJ mol^{-1})	Distance from Best Mode (Å)
1	-29.29	0.00
2	-27.20	1.44
3	-27.20	5.38
4	-26.36	1.46
5	-26.36	5.48

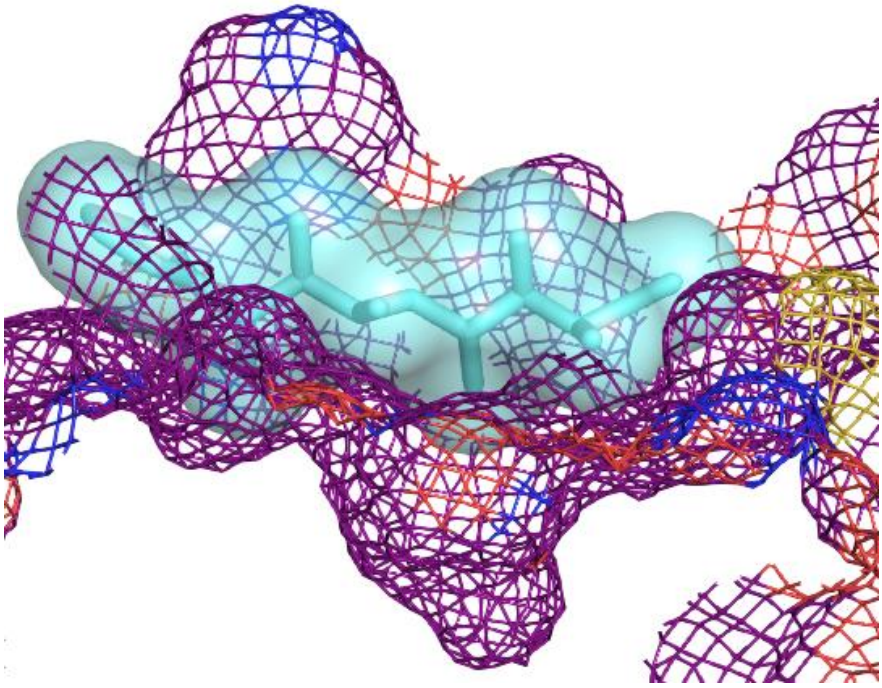


Figure 137. The top ranked binding pose of **111** bound to Mdm2 found by AutoDock Vina. A surface representation of **111** bound to Mdm2 is shown to display the occupancy of the p53 pocket on Mdm2 by **111**.

9.3 Synthetic Procedures

9.3.1 General Synthetic Information

All required fine chemicals were purchased from Acros Organics, Alfa Aesar, Fischer Scientific or Sigma-Aldrich and used as received unless otherwise specified. *s*-butyllithium (*s*-BuLi and *n*-butyllithium (*n*-BuLi) were received from Acros Organics and the molarity was verified by titration with N-benzylbenzamide.[324] (–)-sparteine and (+)-sparteine were distilled over CaH_2 and stored in a young's tube under Argon/N_2 at -5 °C to prevent absorption of atmospheric CO_2. Anhydrous solvents were dried by passing through a modified Grubbs system of alumina columns, manufactured by Anhydrous Engineering, stored over 3Å molecular sieves (25% of total volume) and transferred under N_2 *via* syringe. Ethyl 2,4,6-triisopropylbenzoate (**51**) was provided by Dr. Johan Pradeilles, a previous member of the Aggarwal group.

All air- and water- sensitive reactions were carried out in flame dried glassware under a N_2 atmosphere using standard Schlenk manifold techniques with magnetic stirring. Where reactions were monitored using TLC, aluminium backed plates pre-coated (0.25 mm) with Merck Silica Gel 60 F254 were used and compounds were visualised by exposure to UV light or stained using a 5% solution of phosphomolybdic acid (PMA) in EtOH followed by heating. Where compounds were either purified by flash column chromatography or passed through a plug of silica, Merck Silica Gel 60 (40-63 μm) was used. All mixed solvent eluents are reported as v/v solutions.

^{1}H- and ^{13}C- NMR spectra were recorded on either a JEOL ECZ 300 MHZ, JEOL ECS 400 MHz, Varian 400 MHz, Varian VNMRS 500 MHz NMR spectrometers equipped with direct observe two channel probes, a Bruker AVANCE III HD 500 MHz NMR spectrometer with a 5 mm DCH ^{13}C-^{1}H/D Cryo Probe or a Bruker Avance III HDTM 700 MHz NMR Spectrometer with a 1.7 mm inverse triple resonance micro-Cryo Probe. Spectra were visualised and processed using MestreNova version 14.0. Chemical shifts (δ) are quoted in parts per millions (ppm) and are referenced to the residual proton signals of the solvent. ^{1}H NMR coupling constants are reported in Hz. Data are reported as follows: chemical shift, multiplicity (s = singlet, br. s = broad singlet, d = doublet, t = triplet, q = quartet, quin = quintet, sext = sextet, sept = septet, m = multiplet, dd = doublet of doublets etc.), coupling constant, integration and assignment. Assignment of signals in ^{1}H and ^{13}C NMR spectra was performed using ^{1}H-^{1}H COSY, ^{1}H-^{1}H-gDQFCOSY, DEPT, ^{1}H-^{13}C HSQC, ^{1}H-^{13}C pureshift-HSQC, ^{1}H-^{13}C HMBC, ^{1}H-^{13}C H2BC experiments where appropriate. ^{11}B NMR spectra were measured using Norell S-200-QTZ quartz NMR tubes at 96 or 128 MHz with complete proton decoupling. It should be noted that ^{13}C signals adjacent to boron are generally not observed due to quadrupolar relaxation.

GCMS was performed using an Agilent HP-5MS column (15 m × 0.250 mm), an Agilent 6890 GC, and Agilent 5973 MS system. Compounds were identified through extracted ion chromatogram and molecular ion analysis. Method 70-1X: Inlet temperature 250 °C; Flow rate: 1.0 mL/min; hold at 70 °C for 0 min; ramp 20 °C/min to 200 °C; ramp 45.0 °C/min to 300 °C; hold at 300 °C for 2 min.

Chiral high pressure liquid chromatography (HPLC) separations were performed on an Agilent 1100 Series HPLC unit equipped with UV-vis diode-array detector monitored at 210.8 nm, using Daicel Chiralpak IA, IB or IC columns (4.6 × 250 mm2, 5 μm) fitted with respective guards (4 × 10 mm2).

High resolution mass spectra were recorded by the University of Bristol, School of Chemistry departmental mass spectrometry service using electron spray ionisation (ESI) or matrix assisted laser desorption (MALDI). HRMS ESI was performed on either a Bruker Daltonics Apex IV, 7-Tesla FT-ICR or microTOF II. MALDI was performed in an Applied Biosystems 4700 Proteomics Analyser Instrument. Samples were submitted in either EtOAc or $CHCl_3$.

All Infrared spectra were recorded on the neat compounds using a PerkinElmer Spectrum One FT-IR spectrometer, irradiating between 4000 cm-1 and 600 cm-1. Only strong and selective absorbances above 1400 cm^{-1} are reported. Melting points were measured with a Stuart SMP30 melting point apparatus and are uncorrected. Optical rotations were obtained on a Bellingham + Stanley Ltd. ADP220 polarimeter at 589 nm in a cell with a path length of 1dm. Specific rotations are given in (deg mL)/(g dm). Compound names are those generated by ChemBioDraw 13.0 (PerkinElmer), following the IUPAC nomenclature

9.3.2 General Synthetic Procedures

1. The Homologation of Boronic Esters using Stannanes (GP1)

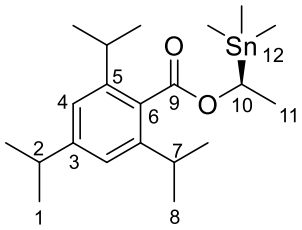
A solution of stannane **53** (1.35 equiv) in a Schlenk reaction vessel was dissolved in anhydrous Et_2O (0.2 M) under an atmosphere of nitrogen. The reaction mixture was cooled to -78 °C and *n*-BuLi (1.5 – 1.6 M in hexanes, 1.30 equiv) was added dropwise to the reaction mixture. The reaction mixture was stirred for 1 h at -78 °C after which the reaction mixture has become a translucent pale yellow solution with no white stannane precipitate remaining. The boronic ester (0.3 M in anhydrous Et_2O, 1 equiv) was added dropwise at -78 °C. The reaction mixture was stirred at -78 °C until the pale yellow solution loses its colour on the addition of the boronic ester. The reaction mixture was then warmed to room temperature, and the 1,2-migration was monitored by ^{11}B NMR. Once the 1,2-migration was complete a white precipitate was formed and filtered through a 1cm plug of silica, to give a pale yellow translucent solution. The silica was washed with Et_2O and the solvent removed (*in situ*) under reduced pressure to give the crude boronic ester. The crude boronic ester was then either purified by flash column chromatography or used directly in subsequent homologations.

2. The Homologation of Boronic Esters using Primary Benzoates (GP2)

A solution of benzoate **121** or **122** (1.35 equiv) and (+)-sparteine (1.30 equiv) in a Schlenk reaction vessel was dissolved in anhydrous Et_2O (0.2 M) under an atmosphere of nitrogen. The reaction mixture was cooled to -78 °C and *s*-BuLi (1.3 - 1.4 M in hexanes, 1.30 equiv) was added dropwise to the reaction mixture. The reaction mixture was stirred for 1 h at -78 °C during which it became dark red in colour. The boronic ester (0.5 M in anhydrous Et_2O, 1 equiv) was added dropwise at -78 °C. The reaction mixture was stirred at -78 °C until the dark red solution loses its colour and becomes yellow. The reaction mixture was then warmed to room temperature, and the 1,2-migration was monitored by ^{11}B NMR. The reaction mixture was then quenched with 2M HCl and the aqueous layer was washed with Et_2O three times. The organic layers were combined, dried using MgSO4, filtered and concentrated in vacuo. The crude boronic ester was then purified by flash column chromatography.

9.3.3 Preparation of Compounds

1-(Trimethylstannyl)ethyl 2,4,6-triisopropylbenzoate, (*S*)-53



An oven dried 1 L, three neck flash, equipped with a 40 mm magnetic stirrer bar was cooled to rt under vacuum. The reaction flask was evacuated and refilled with N_2 (3x). The flask was charged with ethyl 2,4,6-triisopropylbenzoate (7.98 g, 28.9 mmol, 1.0 equiv) and (–)-sparteine (8.3 mL, 37.5 mmol, 1.3 equiv) followed by the addition of anhydrous Et_2O (135 mL) *via* cannula. The solution was cooled to -78 °C and allowed to equilibrate for 10 mins before the addition of *s*BuLi (1.3 mL in hexanes, 28.8 mL, 37.5 mmol, 1.3 equiv) dropwise to the solution. The reaction mixture became dark brown after the addition of *s*-BuLi. The reaction mixture was stirred at -78 °C for 4 h before the addition of Me_3SnCl (1.0 M in hexanes, 37.5 mL, 37.5 mmol, 1.3 equiv) to the reaction mixture. The reaction mixture was stirred at -78 °C for 20 mins, after which it had become a yellow solution. It was then warmed to rt and stirred for 1 h. The reaction mixture was diluted with 2 M HCl (60 mL) and stirred for a further 20 mins. The organic and aqueous layers were separated, and the organic layer was washed with 2 M HCl (4 x 60 mL). The combined aqueous layers were extracted with Et_2O (3 x 60 mL). The combined organic layers were dried ($MgSO_4$), filtered and concentrated *in vacuo* to give the crude stannane **(*S*)-53** (12.3 g, e.r. 91:9) as an off white solid.

The opposite enantiomer **(*R*)-53** was synthesised in identical yields and e.r. by substituting (–)-sparteine for (+)-sparteine. Stannanes used in this project were dried under high vacuum (1 mbar) with stirring overnight.

Recrystallisation of (S)-53

The crude stannane was dissolved in MeOH (3 mL/g) by bringing the solution to reflux. The solution was then allowed to cool to rt. Crystals appeared after 10 min to 5 h depending on the purity of the stannane. The white crystals were filtered and dried under reduced pressure. The recrystallisation was repeated until the e.r. was >99:1. After two recrystallisations **(*S*)-53** was obtained (6.4 g, 52%, e.r. 99.9:0.1) as a colourless solid.

Sparteine Recovery[27]

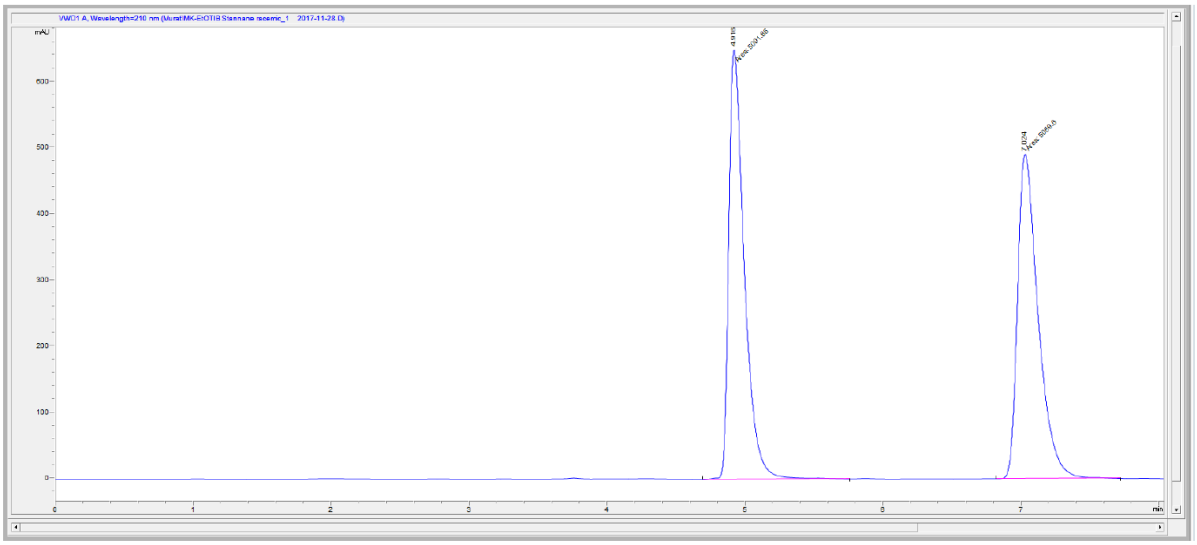
The combined aqueous layers were made basic with NaOH (20%). The aqueous phase was washed with Et_2O (3 x 60 mL) and the combined organic layers were dried (K_2CO_3), filtered and concentrated *in vacuo* to give crude sparteine. Distillation over CaH_2 of the crude material gave (–)-sparteine (6.7 mL, 80%) as a colourless oil.

Spectral data were in accordance with the published values[27]

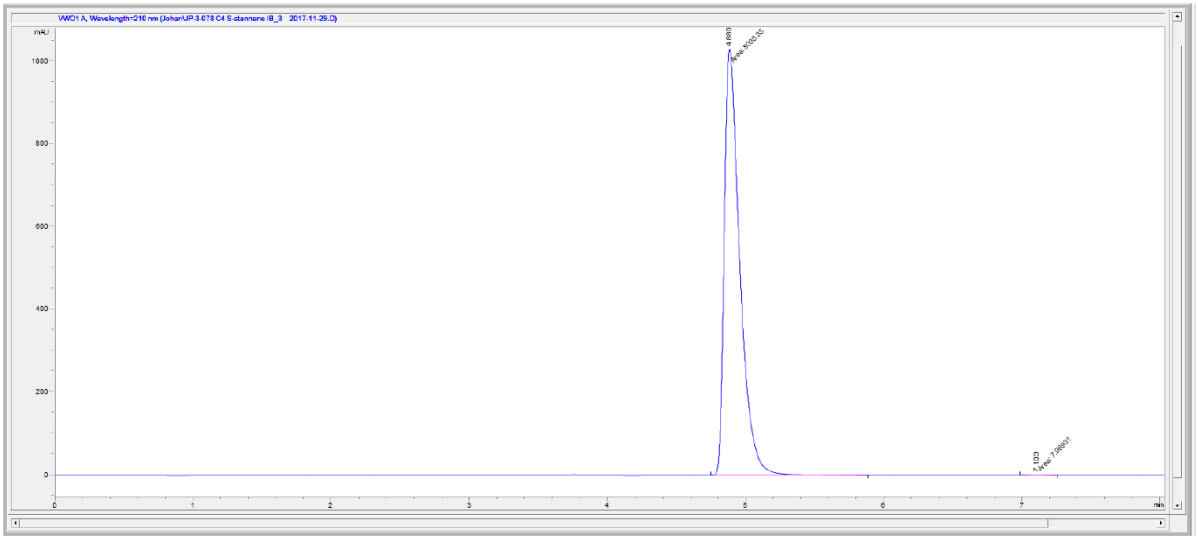
^{1}H NMR (400 MHz, $CDCl_3$) δ: 6.99 (s, 2H, H_4), 5.04 (q, J = 7.4 Hz and quin, J = 7.7 Hz, 1H, H_{10}), 2.79 – 2.92 (m, 3H, H_2 and H_7), 1.59 (d, J = 7.6 Hz and dd, J = 56.5, 7.6 Hz and dd, J = 56.5, 7.6 Hz, 3H, H_{11}), 1.24 (d, J = 6.7 Hz, 18H, H_1 and H_8), 0.18 (s and d, J = 52.7 Hz and d, J = 52.7 Hz, 9H, H_{12})

^{13}C NMR (101 MHz, $CDCl_3$) δ: 171.4 (C_9), 150.0 (Ar), 144.9 (Ar), 130.9 (Ar), 120.9 (C_4), 67.1 (C_{10}), 34.5 (C_2), 31.4 (C_8), 24.4 (C_1 and C_8), 24.2 (C_1 or C_8), 24.0 (C_1 or C_8), 19.3 (C_{11}), -9.8 (d, J = 334 Hz (13C-119Sn) and d, J = 318 Hz (13C-117Sn), C_{12}).

Chiral HPLC: (Daicel Chiralpak-IB column (25 cm) with guard, hexane 0.7 mL/min, rt, 210.8 nm): t_R = 6.4 (*S*), 9.2 (*R*), e.r. = 99.9:0.1



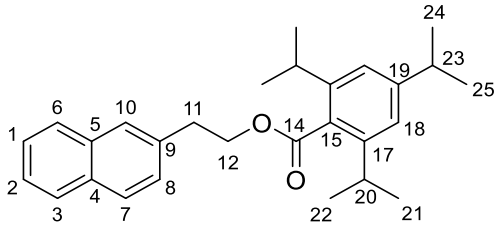
#	Time	Type	Area	Height	Width	Area%	Symmetry
1	4.916	MM	5091.9	649.5	0.1307	50.011	0.516
2	7.024	MM	5089.6	489.9	0.1731	49.989	0.561



#	Time	Type	Area	Height	Width	Area%	Symmetry
1	4.88	MM	8030.4	1028.8	0.1301	99.901	0.507
2	7.103	MM	8	9.3E-1	0.1422	0.099	0.849

Figure 138. The Chiral-HPLC traces for Racemic **53** (top) and **(*S*)-53** (bottom).

2-(naphthalen-2-yl)ethyl 2,4,6-triisopropylbenzoate, 122



To a stirred solution of PPh_3 (1.67 g, 6.39 mmol, 1.1 equiv), 2-naphthalene alcohol (1.00 g, 5.81 mmol, 1 equiv) and 2,4,6-triisopropylbenzoic acid (1.66 g, 6.68 mmol, 1.15 equiv) in anhydrous THF (8 mL) at 0 °C under an atmosphere of N_2, was added DIAD (1.26 mL, 6.39 mmol, 1.1 equiv) dropwise. The reaction mixture was stirred at rt for 4 h, after which the volatiles were removed *in vacuo*. The remaining residue was dissolved in hexane, and the resulting suspension was filtered, and the filter cake washed with hexane. The filtrate was concentrated and the crude product was purified by flash column chromatography (1:10 Et_2O:hexane) to obtain 2.06 g (88%) of **122** as a white solid.

R_f: 0.42 (1:10 Et_2O:hexane)

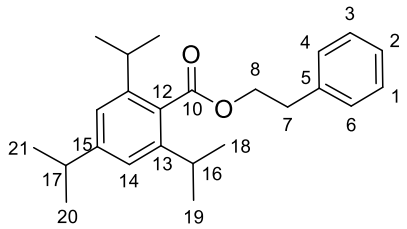
FTIR (ν_{max}/cm^{-1}, neat): 2961, 2930, 2867, 1716, 1606, 1460, 1248, 1082, 746

^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.85-7.76 (m, Naph, 3H), 7.72 (s, 1H, H_{10}), 7.50-7.39 (m, 3H, Naph), 6.97 (s, 2H, H_{18}), 4.68 (t, J = 7.1 Hz, 2H, H_{12}), 3.24 (t, J = 7.1 Hz, 2H, H_{11}), 2.87 (sept, J = 6.7 Hz, 1H, H_{23}), 2.74 (t, J = 6.7 Hz, 2H, H_{20}), 2.13 (d, J = 6.7 Hz, 6H, H_{24}, H_{25}), 1.14 (d, J = 6.7 Hz, 12H, H_{21}, H_{22})

^{13}C NMR (101 MHz, $CDCl_3$) δ: 171.0 (C_{14}), 150.2 (C_{19}), 144.9 (C_{17}), 135.2, 133.7, 132.5, 130.6 (C_{15}), 128.3, 127.7, 127.6, 127.4 (C_{10}), 127.2, 126.2, 125.6, 120.9 (C_{18}), 65.1 (C_{12}), 35.3 (C_{11}), 34.5 (C_{23}), 31.6 (C_{20}), 24.2 (C_{21}, C_{22}), 24.1 (C_{25}, C_{24}).

HRMS: (ESI) calcd. for $C_{28}H_{34}O_2Na$ ($M+Na^+$): 425.2451; Found 425.2446

Phenethyl 2,4,6-triisopropylbenzoate, 121



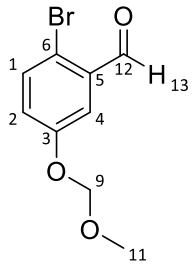
To a stirred solution of PPh_3 (2.89 g, 11 mmol, 1.1 equiv), 2-phenethyl alcohol (1.22 g, 10 mmol, 1 equiv) and 2,4,6-triisopropylbenzoic acid (2.86 g, 11.5 mmol, 1.15 equiv) in anhydrous THF (15 mL) at rt under an atmosphere of N_2, was added DIAD (2.17 mL, 11 mmol, 1.1 equiv) dropwise. The reaction mixture was stirred at 0 °C for 4 h, after which the volatiles were removed *in vacuo*. The remaining residue was dissolved in hexane, and the resulting suspension was filtered, and the filter cake washed with hexane. The filtrate was concentrated and the crude product was purified by flash column chromatography (1:10 Et_2O:hexane) to obtain 3.52 g (93%) of **121** as a white solid.

Spectral data were in accordance with the published values.[325]

^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.37-7.20 (m, 5H, $C_{1,2,3,4,6}$), 7.00 (s, 2H, H_{14}), 4.59 (t, J = 7.1 Hz, 2H, H_8), 3.09 (t, J = 7.1 Hz, 2H, H_7), 2.91 (sept, J = 6.8 Hz, 1H, H_{17}), 2.80 (t, J = 6.8 Hz, 2H, H_{16}), 1.27 (d, J = 6.8 Hz, 6H, H_{20}, H_{21}), 1.21 (d, J = 6.8 Hz, 12H, H_{18}, H_{19}).

^{13}C NMR (101 MHz, $CDCl_3$) δ: 171.0 (C_{10}), 150.2 (C_{15}), 144.9 (C_{13}), 137.8 (C_5), 130.6 (C_{12}), 128.9 ($C_{1,3}$), 128.6 ($C_{4,6}$), 126.7 (C_2), 120.9 (C_{14}), 65.3 (C_8), 35.1 (C_7), 34.6 (C_{17}), 31.6 (C_{16}), 24.2 (C_{18}, C_{19}), 24.1 (C_{20}, C_{21}).

2-Bromo-5-(methoxymethoxy)benzaldehyde, 132



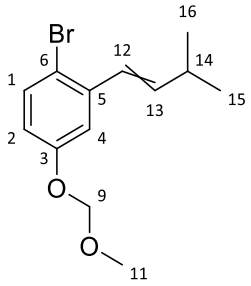
Following a procedure previously reported by Back *et. al.*[263] 2-bromo-5-hydroxybenzaldehyde **131** (5.00 g, 24.87 mmol) was dissolved in 83 mL of anhydrous DCM, and the solution was cooled to 0 °C. Diisopropylethylamine (6.49 mL, 37.31 mmol) was added, followed by methoxymethyl chloride (2.27 mL, 29.85 mmol). The solution was stirred at rt for 24 h and quenched by addition of water (100 mL). The solution was extracted with DCM (3 x 50 mL). The ether extracts were combined, washed with brine (100 mL), dried ($MgSO_4$), and concentrated to afford a brown oil. The crude product was purified by flash column chromatography (2:1 hexane:Et_2O) to afford 5.91 g (96%) of 2-bromo-5-(methoxymethoxy)benzaldehyde **132** as a colourless oil.

Spectral data were in accordance with the published values[263]

^{1}H NMR (400 MHz, CDCl3) δ: 10.29 (s, 1H, H_{13}), 7.56 (d, J = 3.1 Hz, 1H, H_4), 7.53 (d, J = 8.8 Hz, 1H, H_1), 7.14 (dd, J = 8.8, 3.1 Hz, 1H, H_2), 5.19 (s, 2H, H_9), 3.46 (s, 3H, H_{11})

^{13}C NMR (101 MHz, CDCl3) δ: 191.7 (C_{12}), 157.0 (C_3), 134.8 (C_1), 134.3 (C_5), 124.2 (C_2), 118.9 (C_6), 116.6 (C_4), 94.6 (C_9), 56.4 (C_{11})

1-Bromo-4-(methoxymethoxy)-2-(3-methylbut-1-en-1-yl)benzene, 133



Following a procedure previously reported by Spencer *et. al.*[259] *n*-BuLi (1.6 M in hexanes, 14.61 mL, 23.38 mmol) was added to a solution of *iso*butyltriphenylphosphonium bromide (8.97 g, 22.4 mmol) in anhydrous THF (125 mL) at 0 °C dropwise. The reaction mixture became dark red in colour. The reaction mixture was maintained at 0 °C and **132** (5.00 g, 20.33 mmol) was added. The solution was warmed to rt over 1 h after which the solution had turned dark green. The reaction mixture was diluted with hexane (45 mL) and filtered through a silica plug. The solvent was removed *in vacuo* to afford the product as a pale yellow oil. The crude product was purified by flash column chromatography (3:1 hexane:EtOAc) to afford 5.79 g (93%) of 1-bromo-4-(methoxymethoxy)-2-(3-methylbut-1-en-1-yl)benzene **133** as a colourless oil (*E:Z* 1.09:1.0).

R_f: 0.51 (3:1 hexane:EtOAc)

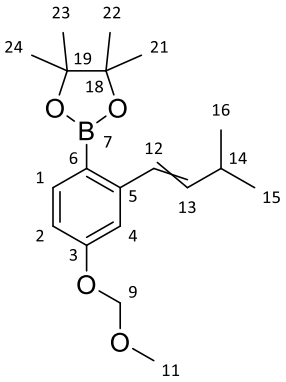
FTIR (ν_{max}/cm^{-1}, neat): 2958, 2930, 2867, 1647, 1462, 1151, 1001

^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.45 (d, J = 8.7 Hz, 1H, H_1), 7.41 (d, J = 8.7 Hz, 1H, H_1), 7.17 (d, J = 3.0 Hz, 1H, H_4), 6.99 (d, J = 3.0 Hz, 1H, H_4), 6.82 (dd, J = 8.7, 3.0 Hz, 1H, H_2), 6.78 (dd, J = 8.7, 3.0 Hz, 1H, H_2) 6.62 (dd, J = 15.9, 1.3 Hz, 1H, H_{12}), 6.27 (d, J = 11.5 Hz, 1H, H_{12}), 6.13 (dd J = 15.9, 6.9, 1H, H_{13}), 5.56 (dd J = 11.5, 10.5, 1H, H_{13}), 5.16 (s, 2H, H_9), 5.15 (s, 2H, H_9), 3.48 (s, 3H, H_{11}), 3.48 (s, 3H, H_{11}), 2.74-2.62 (m, 1H, H_{14}), 2.51 (m, 1H, H_{14}), 1.11 (d, J = 6.8 Hz, 6H, H_{16} and H_{15}), 1.02 (d, J = 6.8 Hz, 6H, H_{16} and H_{15})

^{13}C NMR (101 MHz, $CDCl_3$) δ: 156.7 (C_3), 156.3 (C_3), 141.5 (C_{13}), 141.4 (C_{13}), 139.0 (C_5), 138.7 (C_5), 133.5 (C_1), 133.1 (C_1), 126.1 (C_{12}), 125.9 (C_{12}), 118.6 (C_4), 116.6 (C_2), 116.4 (C_2), 116.00 (C_6), 115.6 (C_6), 114.5 (C_4), 94.8 (C_9), 94.7 (C_9), 56.2 (C_{11}), 56.1 (C_{11}), 31.8 (C_{14}), 27.5 (C_{14}), 23.1 (C_{15}, C_{16}), 22.5 (C_{15}, C_{16})

HRMS: (ESI) calcd. for $C_{13}H_{17}BrO_2$ (M+): 307.0304; Found 307.0309

2-(4-(methoxymethoxy)-2-(3-methylbut-1-en-1-yl)phenyl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane, 134



Following a procedure previously reported by Spencer *et. al.*[259] *n*-BuLi (1.6 M in hexanes, 13.13 mL, 21.04 mmol) was added to a solution of **133** (5.00 g, 17.51 mmol) in 55 mL of anhydrous THF at - 78 °C. The mixture was stirred for 40 minutes after which the solution had turned dark green. 2-Isopropoxy-4,4,5,5-tetramethyl-1,3,2-dioxaborolane (4.30 mL, 21.04 mmol) was added dropwise to the reaction mixture, maintaining the temperature at -78 °C. The reaction mixture turned blue on addition and after 10 minutes the reaction was warmed to rt. Once the solution was at rt a yellow solution was obtained. The reaction was quenched with NH_4Cl (50 mL) and extracted with diethyl ether (3 x 50 mL). The combined organic layers were dried ($MgSO_4$) and concentrated to give a yellow oil. The crude product was purified by flash column chromatography (1:5 EtOAc:hexane) to obtain 5.81 g (95%) of 2-(4-(methoxymethoxy)-2-(3-methylbut-1-en-1-yl)phenyl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane **134** as a colourless oil.

R_f: 0.35 (1:5 EtOAc:hexane)

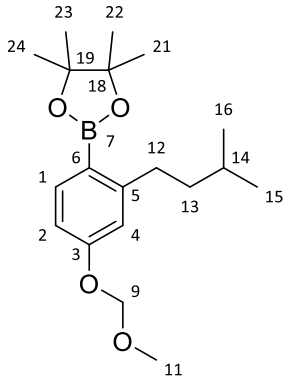
FTIR (ν_{max}/cm^{-1}, neat): 2959, 2931, 2868, 1596, 1343, 1144, 1007

^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.74 (d, J = 8.4 Hz, 1H, H_1), 7.71 (d, J = 8.4 Hz, 1H, H_1), 7.21 (d, J = 2.5 Hz, 1H, H_4), 7.19 (dd, J = 16.0, 1.6 Hz, 1H, H_{12}), 6.97 (d, J = 2.5 Hz, 1H, H_4), 6.89 (dd, J = 8.4, 2.5 Hz, 1H, H_2) 6.87 (dd, J = 8.4, 2.5 Hz, 1H, H_2), 6.78 (d, J = 11.7 Hz, 1H, H_{12}), 6.15 (dd J = 15.7, 6.3, 1H, H_{13}), 5.43 (dd J = 11.7, 10.3, 1H, H_{13}), 5.21 (s, 2H, H_9), 5.19 (s, 2H, H_9), 3.47 (s, 3H, H_{11}), 3.47 (s, 3H, H_{11}), 2.80-2.68 (m, 1H, H_{14}), 2.49 (m, 1H, H_{14}), 1.34 (s, 12H, $C_{21\text{-}24}$), 1.31 (s, 12H, $C_{21\text{-}24}$), 1.11 (d, J = 6.8 Hz, 6H, H_{16} and H_{15}), 1.01 (d, J = 6.8 Hz, 6H, H_{16} and H_{15})

^{13}C NMR (126 MHz, $CDCl_3$) δ: 159.4 (C_3), 158.9 (C_3), 146.4 (C_5), 146.1 (C_5), 139.2 (C_{13}), 138.7 (C_{13}), 137.7 (C_1), 137.4 (C_1), 127.9 (C_{12}), 127.5 (C_4), 116.6 (C_4), 113.7 (C_2), 113.3 (C_2), 111.8 (C_{12}), 94.2 (C_9), 94.0 (C_9), 83.3 (C_{19}), 83.3 (C_{18}), 56.0 (C_{11}), 55.9 (C_{11}), 31.3 (C_{14}), 27.0 (C_{14}), 24.9 ($C_{21\text{-}24}$), 24.8 ($C_{21\text{-}24}$) 23.2 (C_{15}, C_{16}), 22.2 (C_{15}, C_{16})

HRMS: (ESI) calcd. for $C_{19}H_{29}{}^{11}BO_4Na$ ($M+Na^+$): 355.2051; Found 355.2038

2-(2-isopentyl-4-(methoxymethoxy)phenyl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane, 123



To a solution of Pd/C (10 mol%, 0.005 equiv) in EtOH, was added 2-(4-(methoxymethoxy)-2-(3-methylbut-1-en-1-yl)phenyl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane **134** (2.0 g, 6.02 mmol). The reaction was stirred at rt and a balloon of H_2 was secured to a T-connector adaptor that was attached to the reaction vessel. The reaction vessel was evacuated and filled with N_2 (3x) using the T-connector adaptor. This was then repeated, using H_2 in place of N_2. The reaction vessel was left open to H_2 and stirred for 2 h at rt. The reaction mixture was passed through a pad of celite, taking care to not allow it to run dry. The crude NMR revealed complete conversion and the material was used directly in the next step without further purification. 2-(2-isopentyl-4-(methoxymethoxy)phenyl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane, **123**, was obtained in a 98% yield (1.97 g) as a colourless oil.

R_f: 0.41 (1:4 Et_2O:hexane)

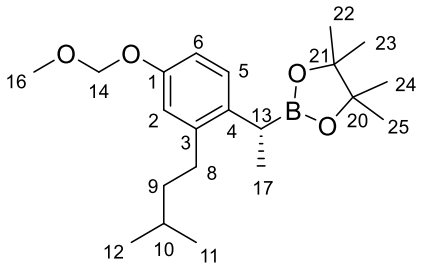
FTIR (ν_{max}/cm^{-1}, neat): 2954, 2931, 2869, 1600, 1345, 1145, 1006

^{1}H NMR (400 MHz, $CDCl_3$) δ: 7.73 (d, J = 8.9 Hz, 1H, H_1), 6.86-6.82 (m, 2H, H_2 and H_4), 5.19 (s, 2H, H_9), 3.47 (s, 3H, H_{11}), 2.86 (m, 2H, H_{12}), 1.65 (sept, J = 6.6 Hz, 1H, H_{14}), 1.45 (m, 2H, H_{13}), 1.33 (s, 12H, $C_{21\text{-}24}$), 0.95 (d, J = 6.6 Hz, 6H, H_{16} and H_{15})

^{13}C NMR (101 MHz, $CDCl_3$) δ: 159.5 (C_3), 153.0 (C_5), 138.1 (C_1), 116.9 (C_4), 112.4 (C_2), 94.1 (C_9), 83.2 (C_{18}, C_{19}), 56.1 (C_{11}), 42.8 (C_{13}), 34.2 (C_{12}), 28.6 (C_{14}), 24.9 ($C_{21\text{-}24}$), 22.7 (C_{15}, C_{16})

HRMS: (ESI) calcd. for $C_{19}H_{31}{}^{11}BO_4Na$ ($M+Na^+$): 357.2208; Found 357.2205

(R)-2-(1-(2-isopentyl-4-(methoxymethoxy)phenyl)ethyl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane, 145



Synthesised according to GP1. Stannane **(*S*)-53** (356 mg, 0.81 mmol, 1.35 eq), *n*-BuLi (1.58 M in hexanes, 0.49 mL, 0.78 mmol, 1.30 eq) and boronic ester **123** (200 mg, 0.6 mmol, 1 eq). Lithiation time of 1 h, borylation time of 30 min, 1,2-migration time of 2 h. The crude boronic ester **145** was not isolated and was used as the crude material in the subsequent homologation. The conversion was monitored by GCMS.

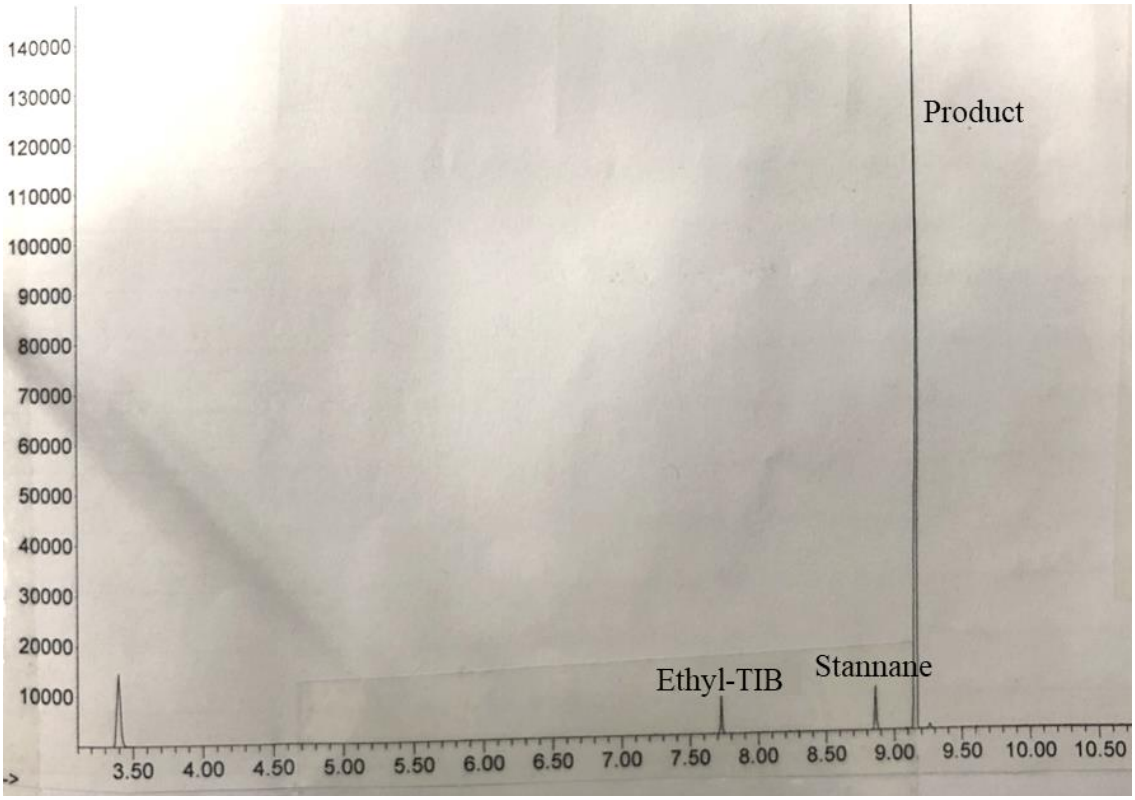
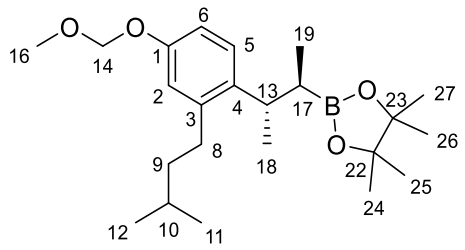


Figure 139. A labelled GCMS trace after the first homologation of boronic ester **145**

2-((2R,3S)-3-(2-isopentyl-4-(methoxymethoxy)phenyl)butan-2-yl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane, 146



Synthesised according to GP1. Stannane **(*S*)-53** (356 mg, 0.81 mmol, 1.35 eq), *n*-BuLi (1.58 M in hexanes, 0.49 mL, 0.78 mmol, 1.30 eq) and boronic ester **145** (217 mg, 0.6 mmol, 1 eq). Lithiation time of 1 h, borylation time of 30 min, 1,2-migration time of 2 h. The crude boronic ester was purified by flash column chromatography (1:10 Et_2O:hexane) to give boronic ester **146** (195 mg, 83% over the two homologations, *d.r.* >99:1 by GCMS) as a colourless oil.

R_f: 0.27 (1:10 Et_2O:hexane)

$[\alpha]^{20}_D$: +16 (*c* 1.0, $CHCl_3$)

FTIR (ν_{max}/cm^{-1}, neat): 2955, 2934, 2870, 1608, 1378, 1315, 1144, 1014

^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.12 (d, *J* = 8.5 Hz, 1H, H_5), 6.86 (dd, *J* = 8.5, 2.8 Hz, 1H, H_6), 6.79 (d, *J* = 2.8 Hz, 1H, H_2), 5.15 (s, 2H, H_{14}), 3.49 (s, 3H, H_{16}), 2.95 (dsept, *J* = 10.6, 6.9 Hz, 1H, H_{13}), 2.67 (ddd, *J* = 13.7, 9.3, 7.4 Hz, 1H, $H_{8'}$), 2.55 (ddd, *J* = 13.7, 9.3, 7.4 Hz, 1H, $H_{8''}$), 1.65 (sept, *J* = 6.7 Hz, 1H, H_{10}), 1.48-1.42 (m, 2H, H_9), 1.28 (s, 12H, H_{24-27}), 1.30-1.26 (m, 1H, H_{17}), 1.19 (d, *J* = 6.9 Hz, 3H, H_{18}), 0.97 (d, *J* = 6.7 Hz, 6H, $H_{11,12}$), 0.76 (d, *J* = 7.4 Hz, 3H, H_{19})

^{13}C NMR (126 MHz, $CDCl_3$) δ: 154.9 (C_1), 142.3 (C_3), 138.9 (C_4), 126.8 (C_5), 116.8 (C_2), 113.5 (C_6), 94.7 (C_{14}), 83.0 (C_{22}, C_{23}), 56.0 (C_{16}), 40.9 (C_9), 36.1 (C_{13}), 31.0 (C_8), 28.2 (C_{10}), 25.0 (C_{24-27}), 24.8 (C_{24-27}), 23.4 (C_{18}), 22.7 ($C_{11/12}$), 22.7 ($C_{11/12}$), 14.6 (C_{19}).

HRMS: (ESI) calcd. for $C_{23}H_{39}{}^{11}BO_4Na$ ($M+Na^+$): 413.2838; Found 413.2820

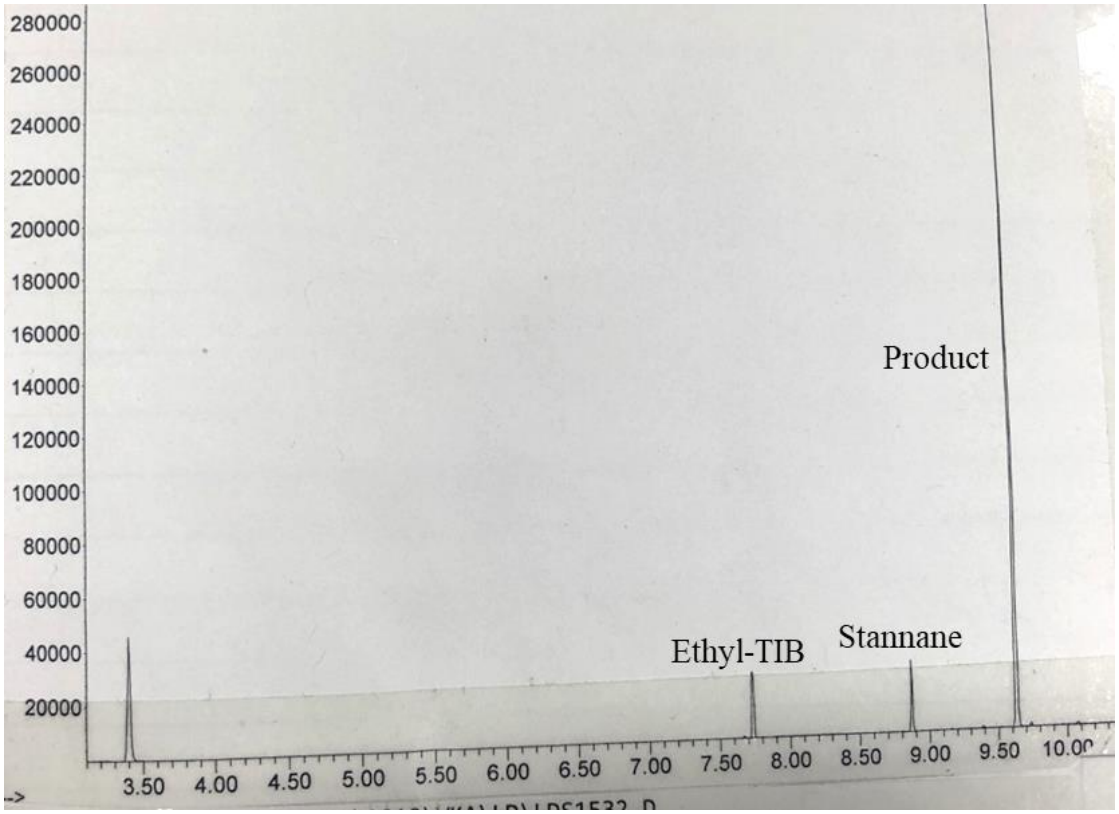
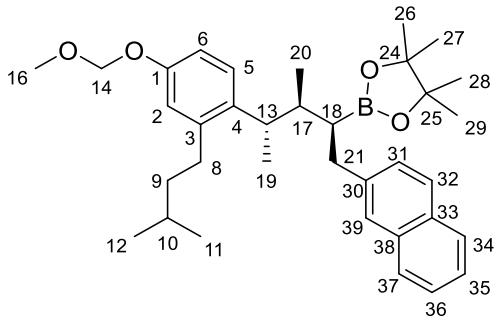


Figure 140. A labelled GCMS trace of the second homologation

2-((2S,3S,4S)-4-(2-isopentyl-4-(methoxymethoxy)phenyl)-3-methyl-1-(naphthalen-2-yl)pentan-2-yl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane, 147



Synthesised according to GP2. Benzoate **122** (288 mg, 0.74 mmol, 1.35 eq), *s*-BuLi (1.3 M in hexanes, 0.57 mL, 0.72 mmol, 1.30 eq), (+)-sparteine (0.17 mL, 0.72 mmol, 1.30 eq) and boronic ester **145** (215 mg, 0.55 mmol, 1 eq). Lithiation time of 1 h, borylation time of 2 h 30 min, 1,2-migration time of 2 h. The crude boronic ester was purified by flash column chromatography (1:10 Et_2O:hexane) to give boronic ester **147** (203 mg, 72%, *d.r.* >95:5 by NMR) as a colourless oil.

R_f: 0.31 (1:10 Et_2O:hexane)

$[\alpha]^{20}_D$: +24 (*c* 1.0, $CHCl_3$)

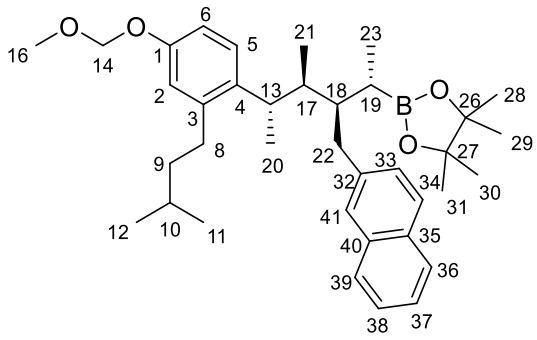
FTIR (ν_{max}/cm^{-1}, neat): 2957, 2931, 2870, 1606, 1496, 1371, 1321, 1145, 1009

^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.87-7.77 (m, 4H, Naph), 7.54-7.40 (m, 3H, Naph), 7.23 (d, *J* = 8.5 Hz, 1H, H_5), 6.93 (dd, *J* = 8.5, 2.9 Hz, 1H, H_6), 6.87 (d, *J* = 2.9 Hz, 1H, H_2), 5.20 (s, 2H, H_{14}), 3.53 (s, 3H, H_{16}), 3.13 (dq, *J* = 10.6, 6.7 Hz, 1H, H_{13}), 3.00 (dd, *J* = 13.9, 11.5 Hz, 1H, $H_{21'}$), 2.94 (dd, *J* = 13.9, 3.9 Hz, 1H, $H_{21''}$), 2.76 (ddd, *J* = 13.6, 9.9, 6.5 Hz, 1H, $H_{8'}$), 2.61 (ddd, *J* = 13.6, 9.9, 6.5 Hz, 1H, $H_{8''}$), 2.08 (dqd, *J* = 9.0, 6.8, 2.8 Hz, 1H, H_{17}), 2.00 (ddd, *J* = 11.5, 3.9, 2.8 Hz, 1H, H_{18}), 1.69 (sept, *J* = 6.6 Hz, 1H, H_{10}), 1.56-1.47 (m, 2H, H_9), 1.38 (d, *J* = 6.7 Hz, 3H, H_{19}), 1.17 (s, 12H, $H_{26\text{-}29}$), 1.04 (d, *J* = 6.6 Hz, 3H, $H_{11/12}$), 1.02 (d, *J* = 6.6 Hz, 3H, $H_{11/12}$), 0.86 (d, *J* = 6.8 Hz, 3H, H_{20})

^{13}C NMR (126 MHz, $CDCl_3$) δ: 154.9 (C_1), 142.0 (C_3), 141.0 (C_{30}), 139.2 (C_4), 133.7 (C_{38}), 132.0 (C_{33}), 128.2 (C_{31}), 127.7, 127.6, 127.5, 127.3 (C_5), 127.1 (C_{39}), 125.8, 124.9, 117.0 (C_2), 114.1 (C_6), 94.7 (C_{14}), 83.2 ($C_{24, C25}$), 56.0 (C_{16}), 40.9 (C_9), 40.6 (C_{17}), 36.4 (C_{13}), 31.5 (C_8), 30.6 (C_{21}), 28.5 (C_{18}), 28.3 (C_{10}), 25.0 ($C_{26\text{-}29}$), 24.6 ($C_{26\text{-}29}$), 22.8 ($C_{11/12}$), 22.7 ($C_{11/12}$), 21.5 (C_{19}), 16.5 (C_{20}).

HRMS: (ESI) calcd. for $C_{35}H_{49}{}^{11}BO_4Na$ ($M+Na^+$): 567.3622; Found 567.3628

2-((2S,3R,4S,5S)-5-(2-isopentyl-4-(methoxymethoxy)phenyl)-4-methyl-3-(naphthalen-2-ylmethyl)hexan-2-yl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane, 148



Synthesised according to GP1. Stannane **(*R*)-53** (221 mg, 0.5 mmol, 1.35 eq), *n*-BuLi (1.58 M in hexanes, 0.31 mL, 0.48 mmol, 1.30 eq) and boronic ester **147** (203 mg, 0.37 mmol, 1 eq). Lithiation time of 1 h, borylation time of 3 h, 1,2-migration time of 2 h. The crude boronic ester was purified by flash column chromatography (1:10 Et_2O:hexane) to give boronic ester **148** (188 mg, 89%, *d.r.* >95:5 by NMR) as a colourless oil.

R_f: 0.17 (1:10 Et_2O:hexane)

$[\alpha]^{20}_D$: +51 (*c* 1.0, $CHCl_3$)

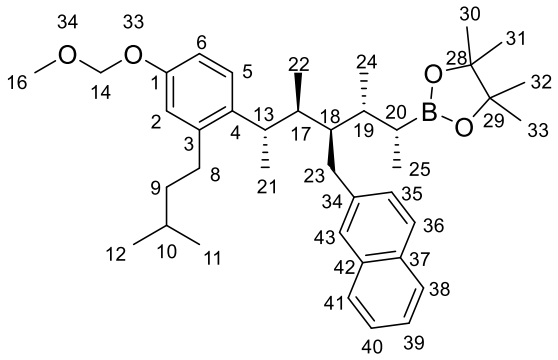
FTIR (ν_{max}/cm^{-1}, neat): 2956, 2870, 1606, 1497, 1463, 1370, 1313, 1144, 1013

^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.83-7.73 (m, 4H, $C_{34, 36, 39, 41}$), 7.51 (dd, *J* = 8.4, 1.4 Hz, 1H, C_{33}), 7.45 (td, *J* = 7.9, 1.4 Hz, 1H, C_{38}), 7.41 (td, *J* = 7.6, 1.4 Hz, 1H, C_{37}), 7.06 (d, *J* = 8.6 Hz, 1H, H_5), 6.84 (dd, *J* = 8.6, 2.3 Hz, 1H, H_6), 6.74 (d, *J* = 2.3 Hz, 1H, H_2), 5.14 (s, 2H, H_{14}), 3.48 (s, 3H, H_{16}), 2.88 (dd, *J* = 14.3, 6.5 Hz, 1H, $H_{22'}$), 2.76-2.66 (m, 2H, $H_{22''}$, H_{13}), 2.56 (br. dddd *J* = 9.6, 6.8, 6.5, 1.9 Hz, 1H, H_{18}), 2.46 (ddd, *J* = 14.1, 9.6, 6.4 Hz, 1H, $H_{8'}$), 2.36 (ddd, *J* = 14.1, 9.6, 6.4 Hz, 1H, $H_{8''}$), 1.96 (br. m, 1H, H_{17}), 1.39 (sept, *J* = 6.8 Hz, 1H, H_{10}), 1.25-1.18 (m, 3H, H_{19}, H_9), 1.18 (s, 6H, $H_{28\text{-}31}$), 1.17 (s, 6H, $H_{28\text{-}31}$), 1.06 (br. d, *J* = 6.9 Hz, 6H, H_{20}, H_{23}), 0.84 (d, *J* = 6.8 Hz, 6H, H_{11}, H_{12}), 0.67 (d, *J* = 6.8 Hz, 3H, H_{21})

^{13}C NMR (126 MHz, $CDCl_3$) δ: 154.6 (C_1), 141.9 (C_3), 140.3 (C_{32}), 139.9 (C_4), 133.8 (C_{40}), 132.2 (C_{35}), 128.2, 127.8 (C_{33}), 127.7, 127.6, 127.4, 127.1 (C_5), 125.9 (C_{38}), 125.1 (C_{37}), 116.7 (C_2), 113.8 (C_6), 94.8 (C_{14}), 83.0 (C_{26}, C_{27}), 56.1 (C_{16}), 41.7 (C_{17}), 41.4 (C_{18}), 40.5 (C_9), 36.5 (C_{13}), 36.1 (C_{22}), 31.4 (C_9), 27.8 (C_{10}), 25.0 ($C_{28\text{-}31}$), 24.9 ($C_{28\text{-}31}$), 22.8 ($C_{11/12}$), 22.7 ($C_{11/12}$), 21.6 (C_{20}), 21.6 (C_{19}), 13.6 (C_{23}), 13.0 (C_{21}).

HRMS: (ESI) calcd. for $C_{37}H_{53}{}^{11}BO_4Na$ ($M+Na^+$): 595.3936; Found 595.3892

2-((2R,3R,4R,5S,6S)-6-(2-isopentyl-4-(methoxymethoxy)phenyl)-3,5-dimethyl-4-(naphthalen-2-ylmethyl)heptan-2-yl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane, 149



Synthesised according to GP1. Stannane **(*S*)-53** (196 mg, 0.45 mmol, 1.35 eq), *n*-BuLi (1.58 M in hexanes, 0.27 mL, 0.43 mmol, 1.30 eq) and boronic ester **148** (188 mg, 0.33 mmol, 1 eq). Lithiation time of 1 h, borylation time of 3 h, 1,2-migration time of 2 h. The crude boronic ester was purified by flash column chromatography (1:10 Et_2O:hexane) to give boronic ester **149** (170 mg, 86%, *d.r.* >95:5 by NMR) as a colourless oil.

R_f: 0.27 (1:10 Et_2O:hexane)

$[\alpha]^{20}_D$: +35 (*c* 1.0, $CHCl_3$)

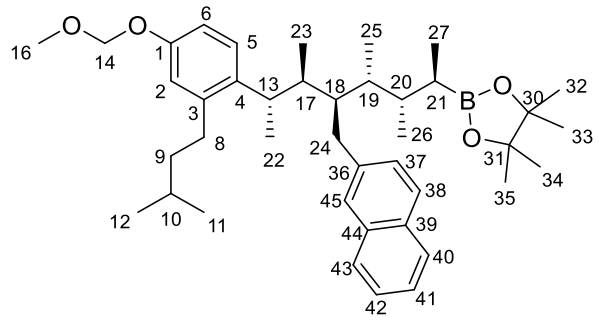
FTIR (ν_{max}/cm^{-1}, neat): 2957, 2872, 1606, 1497, 1461, 1369, 1311, 1147, 1013

^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.83-7.76 (m, 3H, Naph), 7.69 (s, 1H, H_{43}), 7.48-7.39 (m, 3H, Naph), 7.12 (d, J = 8.0 Hz, 1H, H_5), 6.85 (dd, J = 8.0, 2.3 Hz, 1H, H_6), 6.75 (d, J = 2.3 Hz, 1H, H_2), 5.15 (s, 2H, H_{14}), 3.49 (s, 3H, H_{16}), 2.93 (dd, J = 14.3, 5.1 Hz, 1H, $H_{23'}$), 2.81-2.72 (m, 2H, $H_{23''}$, H_{13}), 2.54-2.45 (br. m, 1H, $H_{8'}$), 2.43-2.34 (br. m, 1H, $H_{8''}$), 2.31 (br. dddd, 7.8, 6.6, 5.1, 2.5 Hz, 1H, H_{18}), 1.98 (br. dqd, J = 9.5, 6.7, 2.5 Hz, 1H, H_{17}), 1.87 (sept, J = 6.6 Hz, 1H, H_{19}), 1.40 (sept, J = 6.5 Hz, 1H, H_{10}), 1.34 (m, 1H, H_{20}), 1.24 (m, 2H, H_9), 1.18 (s, 12H, H_{30-33}), 1.15 (d, J = 6.8 Hz, 3H, H_{21}), 0.97 (d, J = 7.1 Hz, 3H, H_{24}), 0.87 (d, J = 7.5 Hz, 3H, H_{25}), 0.85 (d, J = 6.6 Hz, 6H, H_{11}, H_{12}), 0.68 (d, J = 6.7 Hz, 3H, H_{22})

^{13}C NMR (126 MHz, $CDCl_3$) δ: 154.6 (C_1), 141.8 (C_3), 140.7 (C_{34}), 139.6 (C_4), 133.7 (C_{26}), 131.9 (C_{27}), 127.9, 127.7, 127.6, 127.4, 127.1 (C_5), 127.1 (C_{43}), 125.7 (C_{40}), 124.9 (C_{39}), 116.6 (C_2), 113.9 (C_6), 94.6 (C_{14}), 82.7 (C_{28}, C_{29}), 55.9 (C_{16}), 42.6 (C_{18}), 40.4 (C_9), 39.7 (C_{17}), 36.7 (C_{13}), 36.5 (C_{19}), 35.0 (C_{23}), 31.3 (C_8), 27.8 (C_{10}), 24.8 (C_{30-33}), 24.8 (C_{30-33}), 22.6 (C_{11}, C_{12}), 21.8 (C_{21}), 20.4 (C_{20}), 15.3 (C_{24}), 13.9 (C_{22}), 10.9 (C_{25}).

HRMS: (ESI) calcd. for $C_{39}H_{57}{}^{11}BO_4Na$ ($M+Na^+$): 623.4249; Found 623.4239

2-((2R,3S,4R,5R,6S,7S)-7-(2-isopentyl-4-(methoxymethoxy)phenyl)-3,4,6-trimethyl-5-(naphthalen-2-ylmethyl)octan-2-yl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane, 150



Synthesised according to GP1. Stannane **(*S*)-53** (186 mg, 0.38 mmol, 1.35 eq), *n*-BuLi (1.58 M in hexanes, 0.23 mL, 0.36 mmol, 1.30 eq) and boronic ester **149** (170 mg, 0.28 mmol, 1 eq). Lithiation time of 1 h, borylation time of 3 h, 1,2-migration time of 2 h. The crude boronic ester was purified by flash column chromatography (1:10 Et_2O:hexane) to give boronic ester **150** (133 mg, 76%, *d.r.* >95:5 by NMR) as a colourless oil.

R_f: 0.33 (1:10 Et_2O:hexane)

$[\alpha]^{20}_D$: +38 (*c* 1.0, $CHCl_3$)

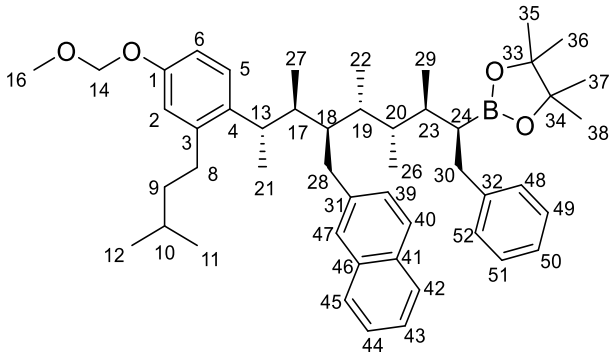
FTIR (ν_{max}/cm^{-1}, neat): 2958, 2871, 1606, 1497, 1463, 1371, 1312, 1146, 1016

^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.83-7.75 (m, 3H, Naph), 7.70 (s, 1H, H_{45}), 7.47-7.38 (m, 3H, Naph), 7.08 (d, J = 8.4 Hz, 1H, H_5), 6.85 (dd, J = 8.4, 2.4 Hz, 1H, H_6), 6.75 (d, J = 2.4 Hz, 1H, H_2), 5.14 (s, 2H, H_{14}), 3.49 (s, 3H, H_{16}), 2.90 (dd, J = 15.0, 5.2 Hz, 1H, $H_{24'}$), 2.81-2.72 (m, 2H, $H_{24''}$, H_{13}), 2.52-2.44 (br. m, 1H, $H_{8'}$), 2.41-2.34 (br. m, 1H, $H_{8''}$), 2.32 (br. dddd, 8.0, 6.9, 5.2, 2.5 Hz, 1H, H_{18}), 1.98 (br. dqd, J = 9.5, 6.9, 2.5 Hz, 1H, H_{17}), 1.79-1.69 (m, 2H, H_{19}, H_{20}), 1.39 (sept, J = 6.6 Hz, 1H, H_{10}), 1.31-1.19 (m, 3H, H_9, H_{21}), 1.13 (d, J = 6.6 Hz, 3H, H_{22}), 1.10 (s, 6H, $H_{32\text{-}35}$), 1.07 (s, 6H, $H_{32\text{-}35}$), 0.92 (d, J = 7.2 Hz, 3H, H_{26}), 0.85 (d, J = 6.9 Hz, 3H, H_{25}), 0.83 (d, J = 6.4 Hz, 6H, H_{11}, H_{12}), 0.74 (d, J = 6.6 Hz, 3H, H_{27}), 0.67 (d, J = 6.9 Hz, 3H, H_{23})

^{13}C NMR (126 MHz, $CDCl_3$) δ: 154.8 (C_1), 141.9 (C_3), 141.0 (C_{36}), 139.6 (C_4), 133.8 (C_{44}), 132.0 (C_{39}), 127.9, 127.8, 127.7, 127.6, 127.3, 127.2, 125.8 (C_{42}), 125.0 (C_{51}), 116.8 (C_2), 114.0 (C_6), 94.8 (C_{14}), 82.7 (C_{30}, C_{31}), 56.1 (C_{16}), 41.8 (C_{18}), 40.6 (C_9), 40.4 (C_{17}), 37.3 (C_{20}), 36.7 (C_{19}), 36.6 (C_{13}), 35.4 (C_{24}), 31.4 (C_8), 27.9 (C_{10}), 24.8 ($C_{32\text{-}35}$), 24.6 ($C_{32\text{-}35}$, C_{21}), 22.7 ($C_{11/12}$), 22.7 ($C_{11/12}$), 21.7 (C_{22}), 14.8 (C_{27}), 14.4 (C_{26}), 13.6 (C_{23}), 12.0 (C_{25}).

HRMS: (ESI) calcd. for $C_{41}H_{61}{}^{11}BO_4Na$ ($M+Na^+$): 651.4562; Found 651.4556

2-((2S,3S,4S,5R,6R,7S,8S)-8-(2-isopentyl-4-(methoxymethoxy)phenyl)-3,4,5,7-tetramethyl-6-(naphthalen-2-ylmethyl)-1-phenylnonan-2-yl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane, 120



Synthesised according to GP2. Benzoate **121** (73 mg, 0.21 mmol, 1.35 eq), *s*-BuLi (1.3 M in hexanes, 0.15 mL, 0.2 mmol, 1.30 eq), (+)-sparteine (0.05 mL, 0.2 mmol, 1.30 eq) and boronic ester **150** (97 mg, 0.15 mmol, 1 eq). Lithiation time of 1 h, borylation time of 3 h, 1,2-migration time of 2 h. The crude boronic ester was purified by flash column chromatography (1:10 Et_2O:hexane) to give boronic ester **120** (92 mg, 84%, *d.r.* >95:5 by NMR) as a colourless oil.

R_f: 0.26 (1:10 Et_2O:hexane)

$[\alpha]^{20}_D$: +24 (*c* 1.0, $CHCl_3$)

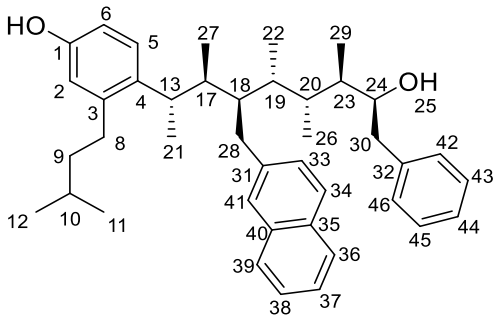
FTIR (ν_{max}/cm^{-1}, neat): 2962, 2874, 1604, 1496, 1465, 1375, 1148, 1018

^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.81-7.71 (m, 4H, Ar), 7.47-7.37 (m, 3H, Ar), 7.13-7.02 (m, 4H, Ar), 6.94 (d, *J* = 8.4 Hz, 2H, H_5, Ar), 6.86 (dd, *J* = 8.4, 2.2 Hz, 1H, H_6), 6.76 (d, *J* = 2.2 Hz, 1H, H_2), 5.14 (s, 2H, H_{14}), 3.48 (s, 3H, H_{16}), 3.01 (dd, *J* = 15.1, 3.6 Hz, 1H, $H_{28'}$), 2.85 (dq, *J* = 10.8, 6.6 Hz, 1H, H_{13}), 2.67 (dd, *J* = 15.1, 6.8 Hz, 1H, $H_{28''}$), 2.57-2.46 (m, 2H, $H_{8'}$, $H_{30'}$), 2.46-2.37 (br. m, 1H, $H_{8''}$), 2.34 (br. dddd, 8.8, 6.8, 3.6, 2.0 Hz, 1H, H_{18}), 2.25 (dd, *J* = 13.3, 3.1 Hz, 1H, $H_{30''}$), 2.03 (br. dqd, *J* = 10.8, 6.7, 2.2 Hz, 1H, H_{17}), 1.85-1.74 (m, 2H, H_{19}, H_{20}), 1.61 (br. dqd, *J* = 9.5, 6.7, 2.5 Hz, 1H, H_{23}), 1.42 (sept, *J* = 7.0 Hz, 1H, H_{10}), 1.38 (ddd, *J* = 12.0, 3.1, 2.5 Hz, 1H, H_{24}), 1.33-1.24 (m, 2H, H_9), 1.18 (d, *J* = 6.6 Hz, 3H, H_{21}), 1.10 (s, 6H, H_{35-38}), 1.09 (s, 6H, H_{35-38}), 0.96 (d, *J* = 6.7 Hz, 3H, H_{29}), 0.83 (m, 9H, H_{11}, H_{12}, H_{22}), 0.68 (d, *J* = 6.7 Hz, 3H, H_{27}), 0.66 (d, *J* = 6.8 Hz, 3H, H_{26}).

^{13}C NMR (126 MHz, $CDCl_3$) δ: 154.8 (C_1), 143.6 (C_{32}), 141.9 (C_3), 141.3 (C_{31}), 139.5 (C_4), 133.8 (C_{46}), 132.0 (C_{41}), 129.1 ($C_{48,52}$), 128.1, 127.9 ($C_{49,51}$), 127.7, 127.7, 127.6, 127.2, 127.0, 126.1, 125.2, 125.1, 116.9 (C_2), 114.1 (C_6), 94.8 (C_{14}), 83.0 (C_{33}, C_{34}), 56.1 (C_{16}), 42.5 (C_{18}), 40.9 (C_{17}), 40.7 (C_9), 36.9, 36.8, 36.2 (C_{13}), 36.0 (C_{28}), 31.5 (C_8), 30.3 (C_{30}), 28.6 (C_{24}), 28.0 (C_{10}), 24.9 (C_{35-38}), 24.7 (C_{35-38}), 22.7 (C_{11}, C_{12}), 21.8 (C_{21}), 16.1 (C_{29}), 12.9 (C_{27}), 12.8 (C_{26}), 12.1 (C_{22}).

HRMS: (MALDI) calcd. for $C_{49}H_{69}{}^{11}BO_4Na$ ($M+Na^+$): 755.5189; Found 755.5180

4-((2S,3S,4R,5R,6R,7R,8S)-8-hydroxy-3,5,6,7-tetramethyl-4-(naphthalen-2-ylmethyl)-9-phenylnonan-2-yl)-3-isopentylphenol, 119



Boronic ester **120** (92 mg, 0.13 mmol) was taken up in MeOH and cooled to 0 °C. Concentrated HCl was added dropwise (5 drops) and the reaction mixture was stirred at 0 °C for 1 h then warmed to rt and stirred overnight. The reaction mixture was diluted with H_2O (5 mL) and Et_2O (5 mL). The phases were separated and the aqueous phase washes with Et_2O (3 x 5 mL). The combined organic phases were dried ($MgSO_4$) and concentrated under reduced pressure. The crude reaction mixture was taken up in THF (2 mL) and cooled to 0 °C. A premixed solution of NaOH (2M):H_2O_2 (30% aq) (2:1 3 mL) was added dropwise at 0 °C, warmed to room temperature and stirred for 4 h. The reaction mixture was diluted with H_2O (5 mL) and Et_2O (5 mL). The phases were separated and the aqueous phase washes with Et_2O (3 x 5 mL). The combined organic phases were washed with H_2O (5 mL), dried ($MgSO_4$) and concentrated under reduced pressure. The crude material was purified by flash column chromatography (7:2:1 hexane:Et_2O:EtOAc) to give **119** as a white solid (62 mg, 86% over two steps, *d.r.* >95:5 by NMR).

m.p.: 161-163 °C (diethyl ether)

R_f: 0.29 (7:2:1 hexane:Et_2O:EtOAc)

$[\alpha]^{20}_D$: +31 (*c* 1.0, $CHCl_3$)

FTIR (ν_{max}/cm^{-1}, neat): 3349, 2963, 2922, 1718, 1601, 1502, 1464, 1371, 1261, 1034, 807

^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.83-7.72 (m, 3H, Ar), 7.69 (s, 1H, H_{41}), 7.46-7.38 (m, 3H, Ar), 7.25-7.15 (m, 3H, Ar), 7.09 (d, *J* = 8.2 Hz, 1H, H_5), 6.87 (m, 2H, Ar), 6.68 (dd, *J* = 8.2, 2.2 Hz, 1H, H_6), 6.60 (d, *J* = 2.2 Hz, 1H, H_2), 4.86 (br.s, 1H, H_{25}), 3.57 (ddd, *J* = 9.9, 4.6, 2.0 Hz, 1H, H_{24}), 3.02 (dd, *J* = 14.9, 3.1 Hz, 1H, $H_{28'}$), 2.91 (br.dq, *J* = 9.5, 6.7 Hz, 1H, H_{13}), 2.67 (dd, *J* = 14.9, 8.3 Hz, 1H, $H_{28''}$), 2.62-2.54 (m, 1H, $H_{8'}$), 2.48-2.40 (br.m, 2H, $H_{8''}$, $H_{30'}$), 2.35 (br.dddd, *J* = 8.3, 8.1, 3.1, 1.7 Hz, 1H, H_{18}), 2.24 (dd, *J* = 13.6, 9.9, 1H, $H_{30''}$), 2.03 (br.dqd, *J* = 9.5, 6.6, 1.7 Hz, 1H, H_{17}), 1.88 (br.dq, *J* = 8.1, 2.4 Hz, 1H, H_{19}), 1.70 (br. dqd, *J* = 8.7, 6.7, 2.4 Hz, 1H, H_{20}), 1.64 (br. dqd, *J* = 8.7, 6.7, 4.6 Hz, 1H, H_{23}), 1.50 (sept,

J = 6.7 Hz, 1H, H_{10}), 1.39-1.31 (m, 2H, H_9), 1.27 (d, J = 6.7 Hz, 3H, H_{21}), 0.93 (d, J = 6.7 Hz, 3H, H_{29}), 0.88 (d, J = 6.9 Hz, 3H, H_{22}), 0.87 (d, J = 6.6 Hz, 6H, H_{11}, H_{12}), 0.70 (d, J = 6.7 Hz, 3H, H_{27}), 0.70 (d, J = 6.7 Hz, 3H, H_{26})

^{13}C NMR (126 MHz, $CDCl_3$) δ: 153.1 (C_1), 142.1 (C_3), 141.0 (C_{31}), 139.4 (C_{32}), 138.0 (C_4), 133.7 (C_{40}), 132.0 (C_{35}), 129.3 ($C_{42,46}$), 128.6 ($C_{43,45}$), 128.1, 127.8, 127.6, 127.5 (C_5), 127.2, 126.4, 126.2, 125.3, 115.8 (C_2), 113.4 (C_6), 74.0 (C_{24}), 42.6 (C_{18}, C_{23}), 40.7 (C_9), 40.3 (C_{17}), 38.3 (C_{30}), 36.5 (C_{13}), 36.3 (C_{20}), 36.0 (C_{19}), 35.8 (C_{28}), 31.3 (C_8), 28.0 (C_{10}), 22.7 ($C_{11/12}$), 22.7 ($C_{11/12}$), 21.7 (C_{21}), 13.4 (C_{27}), 13.1 (C_{26}), 12.1 (C_{22}), 12.0 (C_{29})

HRMS: (MALDI) calcd. for $C_{41}H_{54}O_2Na$ ($M+Na^+$): 601.4016; Found 601.4026

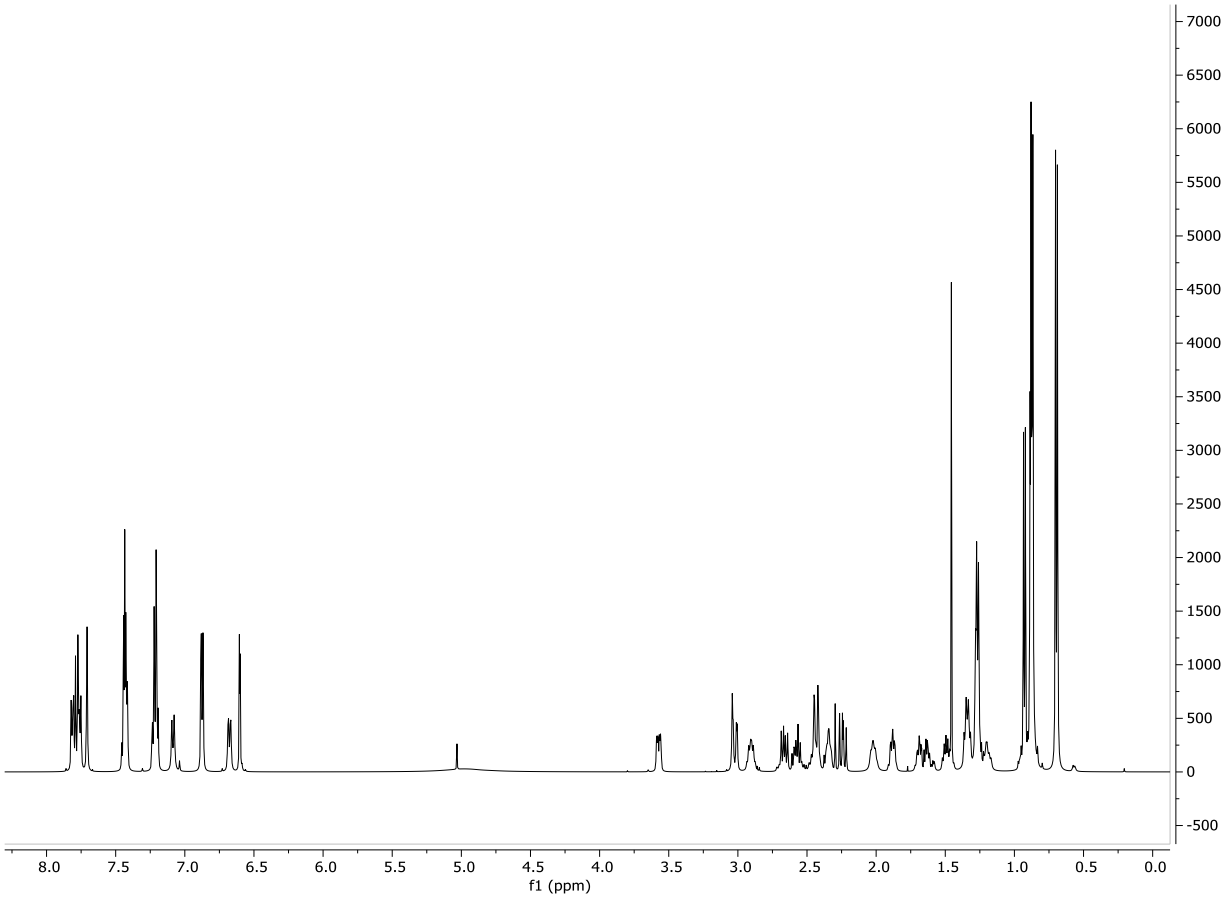


Figure 141. The ^{1}H NMR spectra of **119**

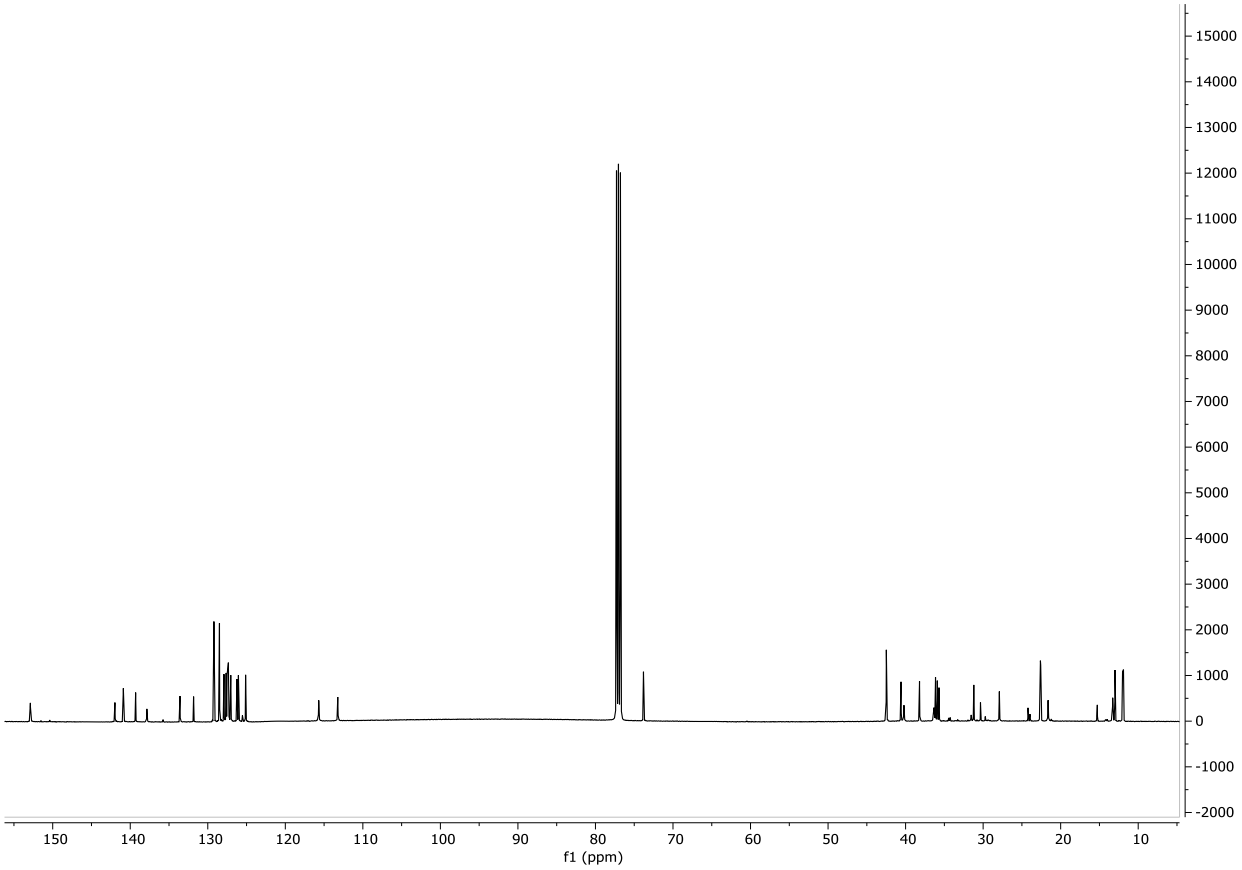
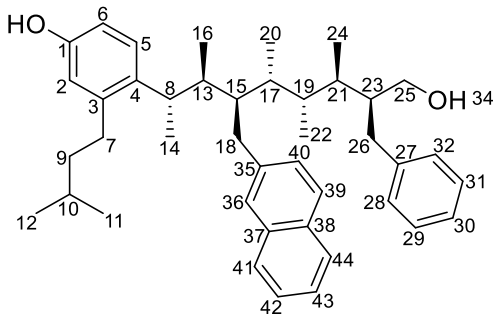


Figure 142. The ^{13}C NMR spectra of **119**

4-((2S,3S,4R,5R,6R,7R,8S)-8-benzyl-9-hydroxy-3,5,6,7-tetramethyl-4-(naphthalen-2-ylmethyl)nonan-2-yl)-3-isopentylphenol, 155



A solution of the boronic ester **120** (95 mg, 0.13 mmol) and bromochloromethane (0.026 mL, 0.39 mmol) in Et_2O (0.65 mL) was cooled to -78 °C. *n*-BuLi (1.6 M in hexanes, 0.21 mL, 0.33 mmol) was added slowly (1 drop every 10 seconds) to the reaction mixture, maintaining a temperature of -78 °C. The reaction mixture was stirred at -78 °C for 1 h and then warmed to rt for 1 h. The reaction mixture was filtered through a plug of wetted (Et_2O) silica of ~1 cm depth to give a colourless solution. The silica was washed with Et_2O (3 mL) and the solvent was removed *in vacuo* to give crude boronic ester. The crude boronic ester was taken up in MeOH and cooled to 0 °C. Concentrated HCl was added (5 drops) and the reaction mixture was stirred at 0 °Cf for 1 h and then warmed to rt overnight. The reaction mixture was diluted with H_2O (5 mL) and Et_2O (5 mL). The phases were separated and the aqueous phase washes with Et_2O (3 x 5 mL). The combined organic phases were dried ($MgSO_4$) and concentrated under reduced pressure. The crude reaction mixture was taken up in THF (2 mL) and cooled to 0 °C. A premixed solution of NaOH (2M):H_2O_2 (30% aq) (2:1 3 mL) was added dropwise at 0 °C, warmed to room temperature and stirred for 4 h. The reaction mixture was diluted with H_2O (5 mL) and Et_2O (5 mL). The phases were separated and the aqueous phase washes with Et_2O (3 x 5 mL). The combined organic phases were washed with H_2O (5 mL), dried ($MgSO_4$) and concentrated under reduced pressure. The crude material was purified by flash column chromatography (7:2:1 hexane:Et_2O:EtOAc) to give **155** as a white solid (59 mg, 77% over three steps, *d.r.* >95:5 by NMR).

m.p.: 169-174 °C (diethyl ether)

R_f: 0.25 (7:2:1 hexane:Et_2O:EtOAc)

$[\alpha]^{20}_D$: +20 (*c* 1.0, $CHCl_3$)

FTIR (ν_{max}/cm^{-1}, neat): 3360, 2959, 2934, 1605, 1502, 1470, 1385, 1240, 1027, 818

^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.80-7.74 (m, 2H, Ar), 7.73-7.67 (m, 2H, Ar), 7.44-7.35 (m, 3H, Ar), 7.15-7.09 (m, 3H, $H_{29,30,31}$), 7.07 (d, *J* = 8.5 Hz, 1H, H_5), 6.85-6.80 (m, 2H, $H_{28,32}$), 6.66 (dd, *J* = 8.5,

2.9 Hz, 1H, H_6), 6.58 (d, *J* = 2.9 Hz, 1H, H_2), 4.62 (br.s, 1H, H_{34}), 3.46 (dd, *J* = 11.1, 4.3 Hz, 1H, $H_{25'}$), 3.37 (dd, *J* = 11.1, 8.9 Hz, 1H, $H_{25''}$), 3.02 (dd, *J* = 15.1, 3.9 Hz, 1H, $H_{18'}$), 2.86 (br. dq, *J* = 9.9, 6.6 Hz, 1H, H_8), 2.63 (dd, *J* = 15.1, 7.2 Hz, 1H, $H_{18''}$), 2.56-2.48 (m, 1H, $H_{7'}$), 2.44-2.37 (br.m, 1H, $H_{7''}$), 2.37 (dd, *J* = 14.1, 2.6 Hz, 1H, $H_{26'}$), 2.36-2.30 (m, 1H, H_{15}), 2.05-1.97 (m, 1H, H_{13}), 2.00 (dd, *J* = 14.1, 11.2 Hz, 1H, $H_{26''}$), 1.84 (br. dqd, *J* = 8.7, 6.9, 2.0 Hz, 1H, H_{17}), 1.82-1.75 (m, 1H, H_{23}), 1.73-1.65 (m, 2H, $H_{19,21}$), 1.44 (sept, *J* = 6.9 Hz, 1H, H_{10}), 1.32-1.26 (m, 2H, H_9), 1.22 (d, *J* = 6.6 Hz, 3H, H_{14}), 0.86 (d, *J* = 6.9 Hz, 3H, H_{20}), 0.83 (d, *J* = 6.5, 9H, $H_{11,12,24}$), 0.68-0.64 (m, 6H, $H_{16,22}$)

^{13}C NMR (126 MHz, $CDCl_3$) δ: 153.0 (C_1), 142.1 (C_3), 141.6 (C_{27}), 141.1 (C_{35}), 138.1, 133.7, 132.0, 129.0 ($C_{28,32}$), 128.4 ($C_{29,31}$), 128.1, 127.7, 127.7, 127.5, 127.0, 126.1, 125.8 (C_{30}), 125.2, 115.8 (C_2), 113.3 (C_6), 64.5 (C_{25}), 44.1 (C_{15}), 42.3 (C_{15}), 40.7 (C_{13}), 40.7 (C_9), 36.9 (C_{17}), 36.2 (C_8), 36.1 (C_{18}), 36.0 ($C_{19,21}$), 32.5 (C_{26}), 31.3 (C_7), 28.0 (C_{10}), 22.7 ($C_{11/12}$), 22.7 ($C_{11/12}$), 21.7 (C_{14}), 13.0 (C_{16}), 12.9 (C_{22}), 12.3 (C_{24}), 12.0 (C_{20}).

HRMS: (ESI) calcd. for $C_{42}H_{56}O_2Na$ ($M+Na^+$): 615.4173; Found 615.4156

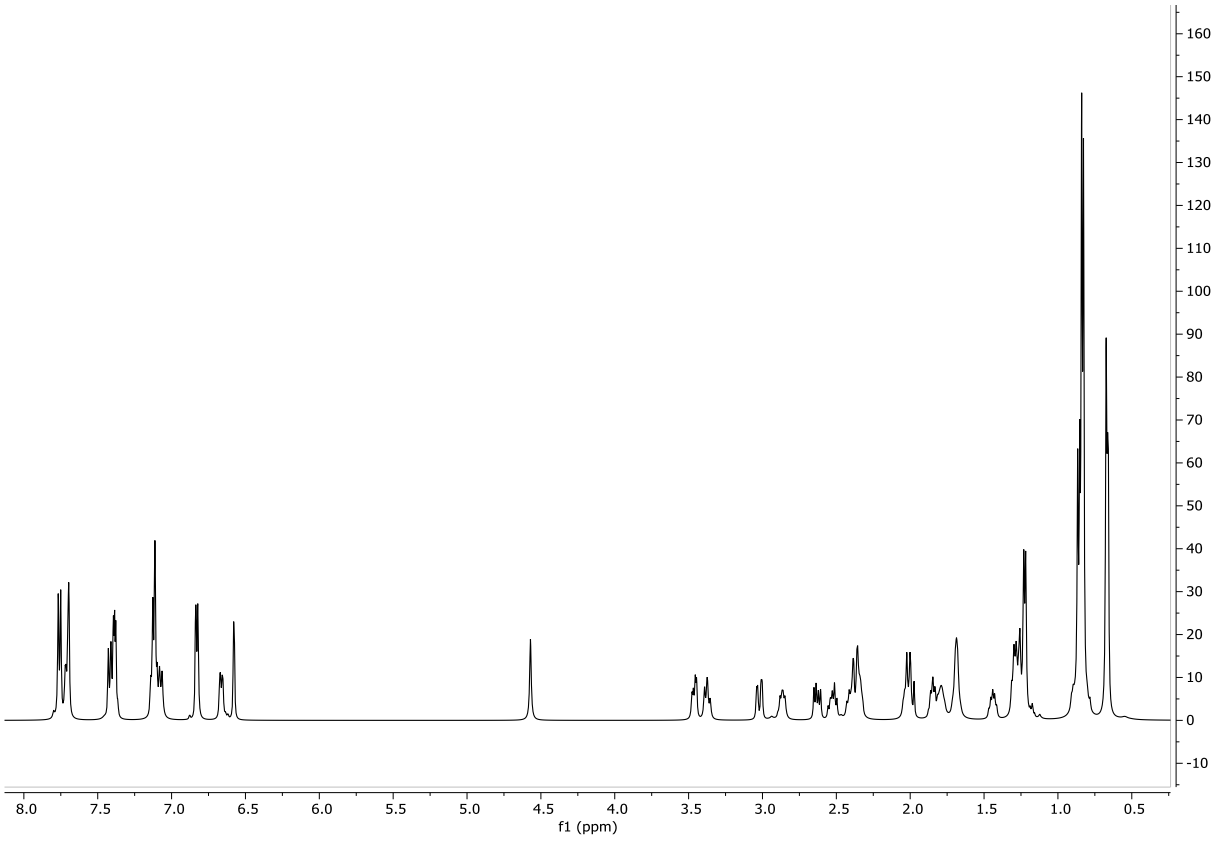


Figure 143. The ^{1}H NMR spectra of **155**

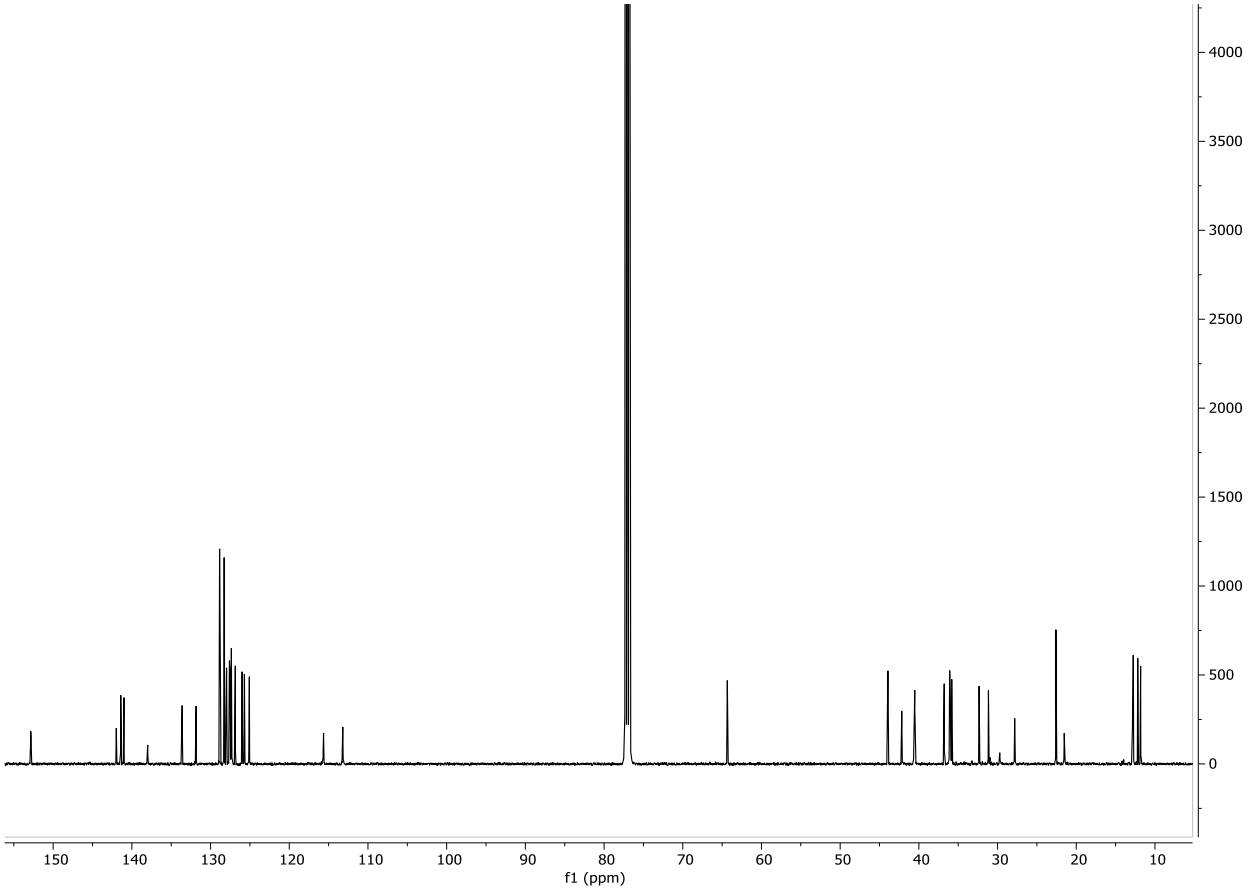
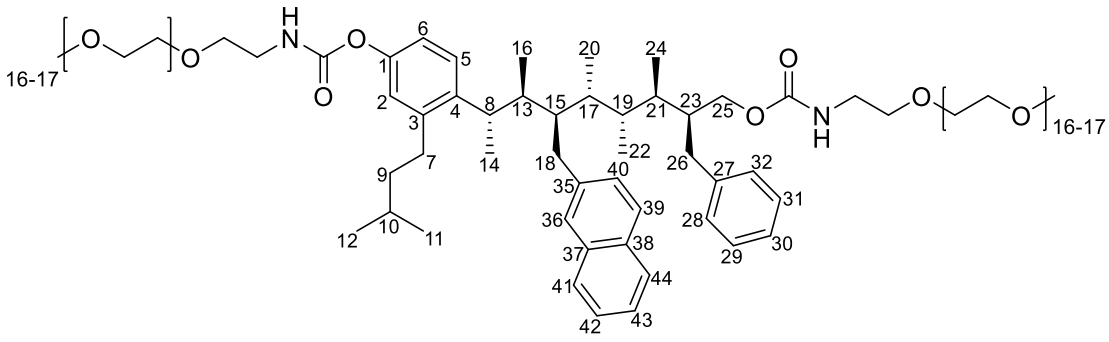


Figure 144. The ^{13}C NMR spectra of **155**

4-((12S,13R,14R,15R,16R,17S,18S)-12-benzyl-13,14,15,17-tetramethyl-16-(naphthalen-2-ylmethyl)-9-oxo-2,5,10-trioxa-8-azanonadecan-18-yl)-3-isopentylphenyl (2-(2-methoxyethoxy)ethyl)carbamate, 157



The bisalcohol **155** (19.0 mg, 0.033 mmol, 1.0 eq) was solubilised in DCM (1 mL) and cooled to 0 °C. Pyridine (0.045 mL, 0.66 mmol, 20.0 eq) was added dropwise followed by the dropwise addition of phosgene (15 wt.% in toluene) (0.49 mL, 0.66 mmol, 20.0 eq). The reaction was warmed to room temperature and stirred for 12 h. The reaction was monitored by TLC (1:10 MeOH:DCM) and if not complete, stirred at 40 °C for 12 h. Upon completion of the reaction, the solvent was removed *in. vacuo* and the crude reaction mixture was resuspended in DCM (1 mL). The reaction mixture was cooled to 0 °C and pyridine was added dropwise (0.023 mL, 0.33 mmol, 10.0 eq). A solution of methoxypolyethylene glycol amine (750 average Mw) in DCM (247 mg, 0.33 mmol, 10 eq) was added dropwise to the reaction mixture. The reaction was warmed to rt and stirred for 12 h. The reaction was monitored by TLC and if not complete, stirred at 40 °C for 12 h. The crude reaction mixture was filtered through a plug of silica (1cm depth, wetted with DCM, eluted with 10% MeOH in DCM). The crude reaction mixture was then purified by reversed phase HPLC (1:10 AcCN:H_2O) to yield the bisPEGylated product **157** (3.7 mg, 11%) as a colourless oil.

R_f: 0.09 (1:10 MeOH:DCM)

$[\alpha]^{20}_D$: +0.42 (*c* 1.0, $CHCl_3$)

FTIR (ν_{max}/cm^{-1}, neat): 3318, 2964, 1731, 1679, 1611, 1465, 1306, 1135, 814

^{1}H NMR (500 MHz, MeOD) δ: 7.85-7.68 (m, 4H, Ar), 7.46 (dd, *J* = 8.4, 0.9 Hz, 1H, Ar), 7.41-7.35 (m, 2H, Ar), 7.13-7.01 (m, 4H, Ar), 6.78 (dd, *J* = 7.7, 1.4 Hz, 2H, Ar), 6.64 (dd, *J* = 8.4, 2.3 Hz, 1H, H_6), 6.54 (d, *J* = 2.3 Hz, 1H, H_2), 3.94 (dd, *J* = 11.2, 8.3 Hz, 1H, $H_{25'}$), 3.88 (dd, *J* = 11.2, 5.2 Hz, 1H, $H_{25''}$), 3.79 (m, 2H, CH_2-PEG), 3.73-3.60 (m, 140H, PEG), 3.59-3.49 (m, 8H, PEG), 3.37 (s, 6H, CH_3-PEG), 3.08 (dd, *J* = 15.1, 3.0 Hz, 1H, $H_{18'}$), 2.92 (br. dq, *J* = 10.1, 6.5 Hz, 1H, H_8), 2.65 (dd, *J* = 15.1, 7.6 Hz, 1H, $H_{18''}$), 2.61-2.53 (m, 1H, $H_{7'}$), 2.43-2.33 (m, 3H, $H_{7''}$, H_{15}, $H_{26'}$), 2.12-2.02 (m, 2H, $H_{26''}$, H_{13}), 1.99-1.94 (br.m, 1H, H_{23}), 1.94-1.81 (br.m, 1H, H_{17}), 1.69 (br.dqd, *J* = 9.7, 6.6, 2.4 Hz, 1H, H_{19}), 1.57 (br.dqd,

J = 9.7, 6.8, 2.8 Hz, 1H, H_{21}), 1.43 (sept, *J* = 6.5 Hz, 1H, H_{10}), 1.36-1.22 (m, 5H, H_9, H_{14}), 0.94-0.85 (br.d, 6H, H_{20}, H_{24}), 0.84 (br.d, *J* = 6.5 Hz, 6H, H_{11}, H_{12}), 0.69 (d, *J* = 6.9 Hz, 3H, H_{16}), 0.65 (d, *J* = 6.6 Hz, 3H, H_{22})

^{13}C NMR (126 MHz, MeOD) δ: 171.5, 164.1, 155.7, 142.6, 142.3, 142.3, 137.6, 135.1, 133.4, 129.9, 129.2, 129.1, 128.7, 128.5, 128.4, 128.1, 127.1, 126.8, 126.2, 116.8 (C_2), 114.4 (C_6), 73.7 (PEG), 73.0 (PEG), 71.6 (PEG), 71.5 (PEG), 71.5 (PEG), 71.4 (PEG), 71.3 (PEG), 67.5 (C_{25}), 62.2, 59.9 (CH_3-PEG), 59.1, 54.2, 43.5 (C_{15}), 42.1 (C_9), 41.8 (C_{13}), 38.0 (C_{17}), 37.9 (C_{19}), 37.3 (C_{18}), 37.2 (C_8), 33.3 (C_{26}), 32.5 (C_7), 29.0 (C_{10}), 23.0 ($C_{11/12}$), 22.9 ($C_{11/12}$), 21.9 (C_{14}), 13.2 (C_{22}), 13.2 (C_{16}), 12.9 (C_{20} or C_{24}), 12.1 (C_{20} or C_{24})

MALDI MS:

Table 37. The Mw distribution table obtained from MALDI for PEG derivative **157**

		[M + Na]$^+$		
n	**Calculated Mw**	**Calculated**	**Observed**	**%Intensity**
25	1895.2	1918.2	1918.2	14
26	1939.2	1962.2	1962.3	21
27	1983.2	2006.2	2006.3	36
28	2027.3	2050.3	2050.3	47
29	2071.3	2094.3	2094.3	63
30	2115.3	2138.3	2138.4	77
31	2159.3	2182.3	2182.4	86
32	2203.4	2226.4	2226.4	92
33	**2247.4**	**2270.4**	**2270.4**	**100**
34	2291.4	2314.4	2314.5	90
35	2335.4	2358.4	2358.5	88
36	2379.5	2402.5	2402.5	81
37	2423.5	2446.5	2446.5	61
38	2467.5	2490.5	2490.6	53
39	2511.5	2534.5	2534.6	40
40	2555.6	2578.6	2578.6	34
41	2599.6	2622.6	2622.6	31
42	2643.6	2666.6	2667.7	26
43	2687.7	2710.7	2710.7	25
44	2731.7	2754.7	2754.7	22
45	2775.7	2798.5	2798.7	19
46	2819.7	2842.7	2842.7	18
47	2863.8	2886.8	2886.8	16
48	2907.8	2930.8	2930.8	14
49	2951.8	2974.8	2974.9	13
50	2995.8	3018.8	3018.9	10

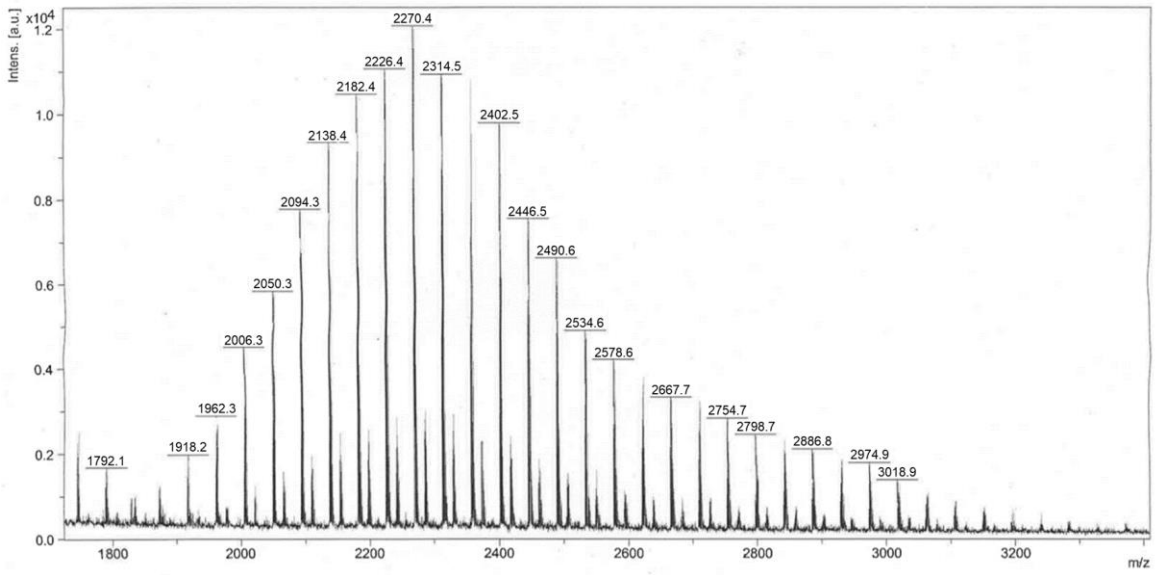


Figure 145. The MALDI MS spectrum of **157**

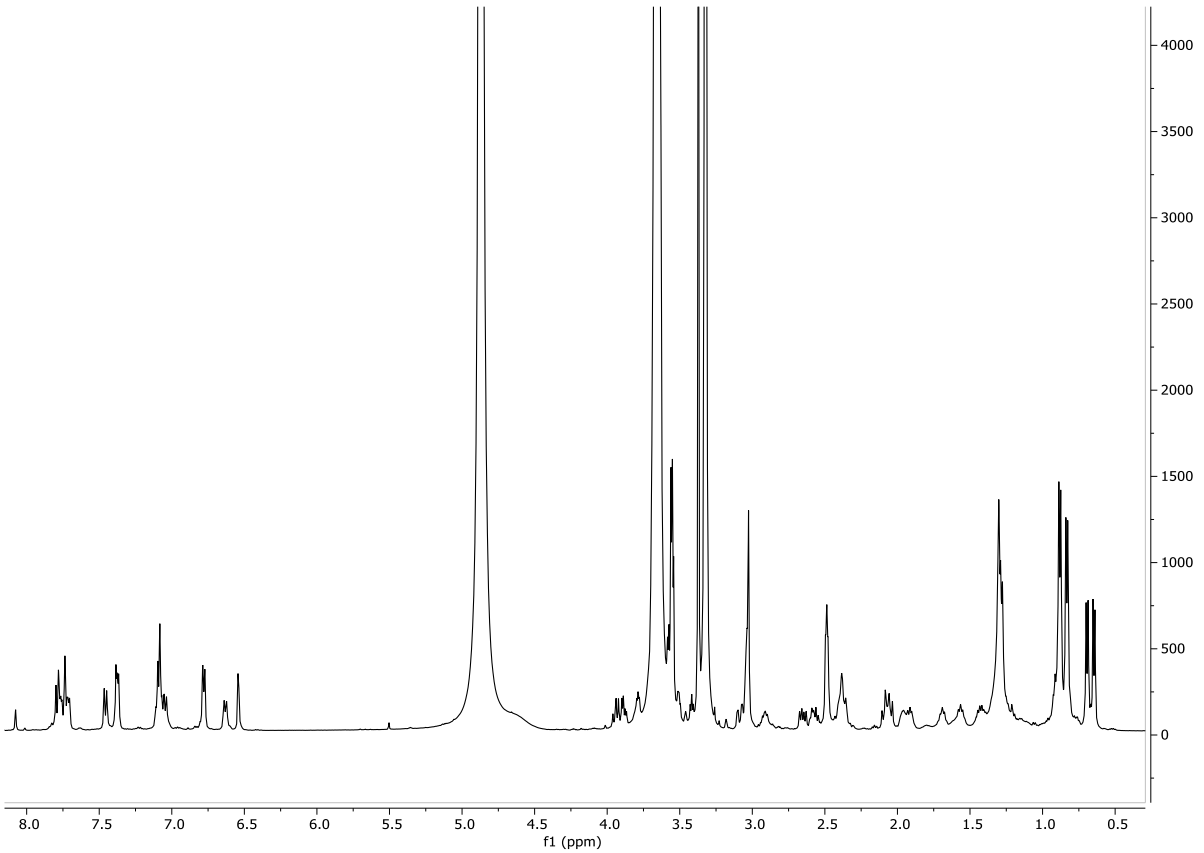


Figure 146. The ^{1}H NMR spectra of **157**

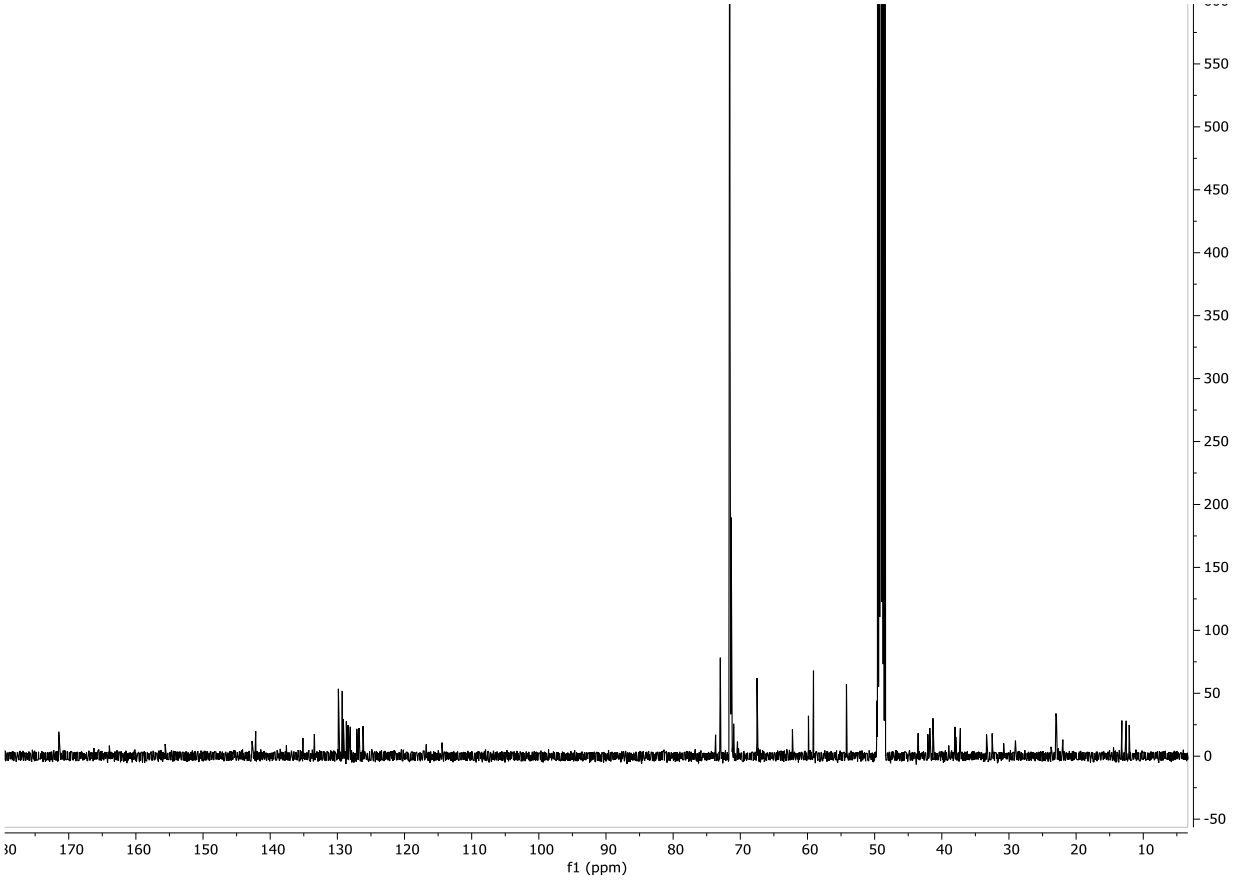
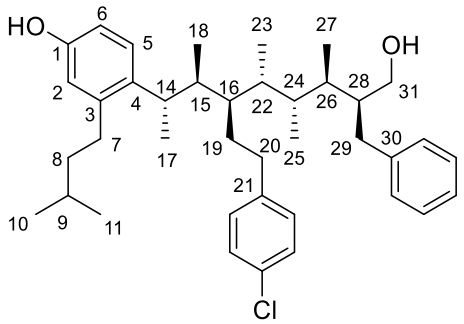


Figure 147. The ^{13}C NMR spectra of **157**

4-((2S,3S,4R,5R,6R,7R,8S)-8-benzyl-4-(4-chlorophenethyl)-9-hydroxy-3,5,6,7-tetramethylnonan-2-yl)-3-isopentylphenol, 163



*Compound **163** was synthesised by Dr. Madhavachary Rudrakshula using a lithiation–borylation sequence resembling **155**.*

R_f: 0.22 (3:7 EtOAc:hexane)

$[\alpha]^{20}_D$: +0.23 (*c* 1.0, $CHCl_3$)

FTIR (ν_{max}/cm^{-1}, neat): 3440, 3158, 2968, 2935, 1605, 1515, 1497, 1382, 1255, 1061

^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.32-7.26 (m, 2H, Ar), 7.23-7.17 (m, 5H, Ar), 7.13-7.08 (m, 2H, Ar), 7.03 (d, *J* = 8.4 Hz, 1H, H_5), 6.66 (dd, *J* = 8.4, 2.7 Hz, 1H, H_6), 6.60 (d, *J* = 2.7 Hz, 1H, H_2), 4.59 (s, 1H,), 3.61 (dd, *J* = 11.0, 4.6 Hz, 1H, $H_{31'}$), 3.57-3.49 (m, 1H, $H_{31''}$), 2.91-2.80 (m, 1H, H_{14}), 2.73 (dd, *J* = 13.5, 2.9 Hz, 1H, $H_{29'}$), 2.71-2.60 (m, 3H), 2.55-2.45 (m, 1H, $H_{7'}$), 2.23 (dd, *J* = 13.5, 11.2 Hz, 1H, $H_{29''}$), 2.19-2.11 (m, 1H, H_{28}), 1.94-1.86 (m, 1H, H_{15}), 1.86-1.79 (m, 1H, H_{26}), 1.79-1.69 (m, 4H, $H_{19',24,16,22}$), 1.63 (sept, *J* = 6.6 Hz, 1H, H_9), 1.48-1.36 (m, 3H, $H_{8,19''}$), 1.17 (d, *J* = 6.7 Hz, 3H, H_{17}), 1.07-1.00 (brs, 1H, H_{32}), 0.97 (d, *J* = 6.5 Hz, 3H, H_{25}), 0.94 (d, *J* = 6.6 Hz, 3H, H_{11}), 0.93 (d, *J* = 6.6 Hz, 3H, H_{10}), 0.89 (d, *J* = 6.8 Hz, 3H, H_{27}), 0.80 (d, *J* = 5.8 Hz, 3H, H_{23}), 0.50 (d, *J* = 6.8 Hz, 3H, H_{18})

^{13}C NMR (126 MHz, $CDCl_3$) δ: 152.9 (C_1), 141.9 (C_3), 141.5 (C_{21}), 141.4 (C_{30}), 137.9 (C_4), 131.5, 129.5, 128.9, 128.6, 128.5, 127.3 (C_5), 126.1, 115.7 (C_2), 113.3 (C_6), 64.5 (C_{31}), 44.0 (C_{28}), 40.8 (C_8), 40.5 (C_{15}), 40.2 (C_{16}), 36.6 (C_{20}), 36.3 (C_{22}), 36.0 (C_{14}), 35.8 (C_{24}), 35.6 (C_{26}), 32.4 (C_{29}), 31.6 (C_{19}), 31.4 (C_7), 28.3 (C_9), 22.7 (C_{11}), 22.6 (C_{10}), 21.5 (C_{17}), 12.6 (C_{25}), 12.3 (C_{27}), 12.2 (C_{18}), 11.8 (C_{23})

HRMS: (MALDI) calcd. for $C_{39}H_{55}ClO_2Na$ ($M+Na^+$): 613.3788; Found 613.3796

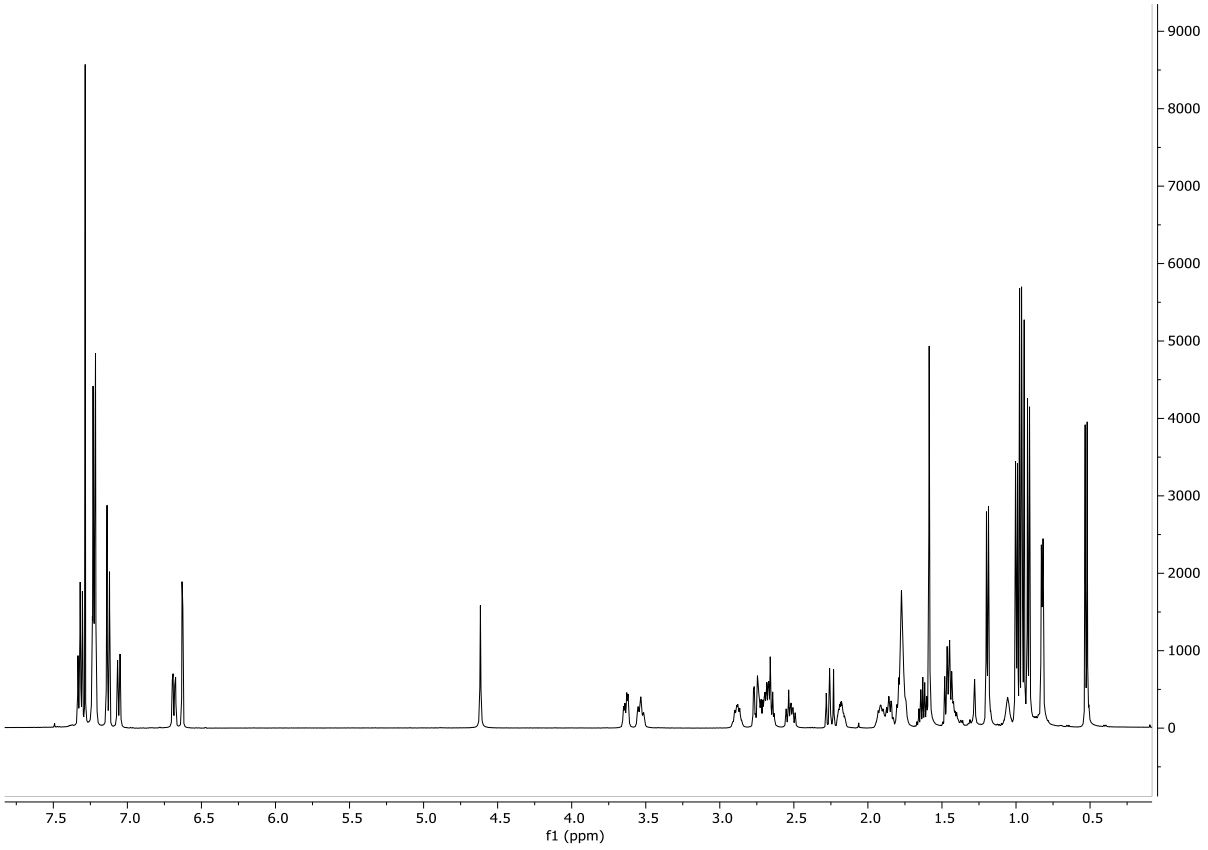


Figure 148. The ^{1}H NMR spectra of **163**

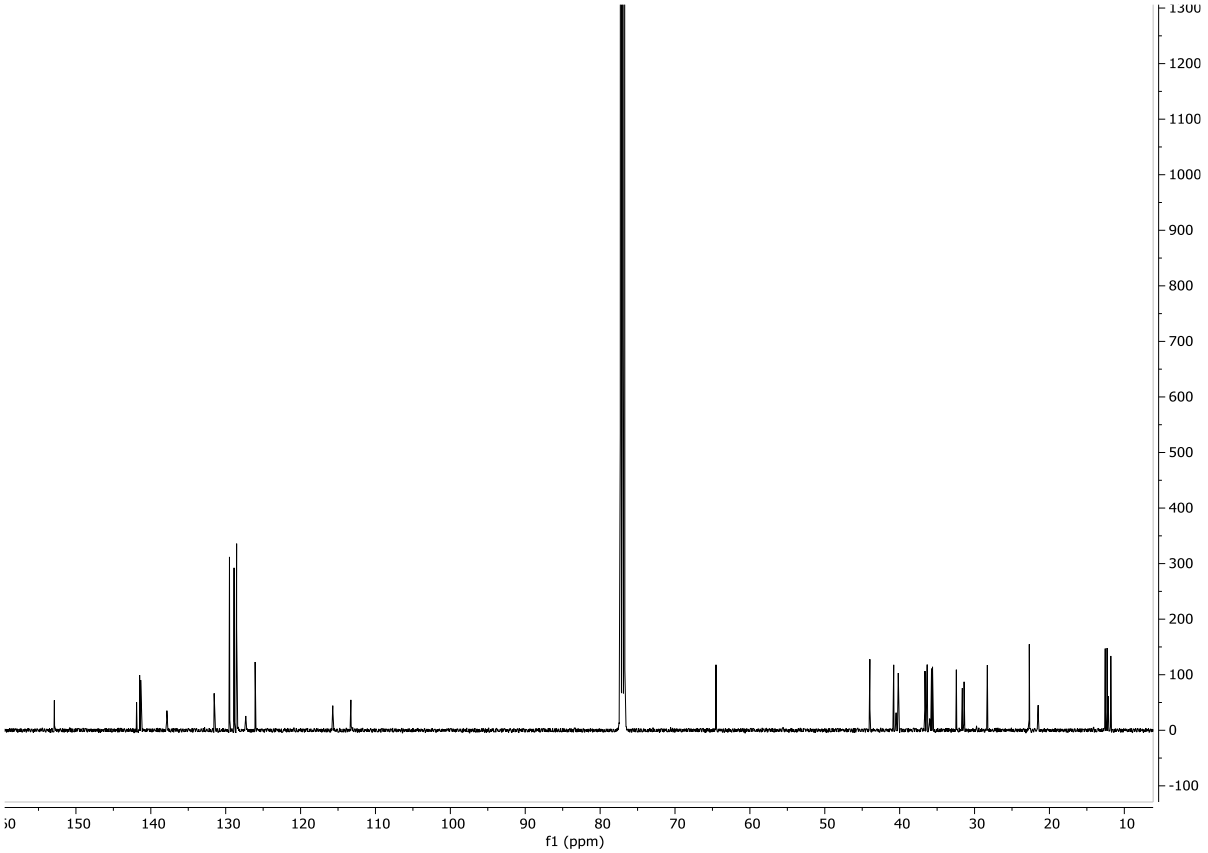
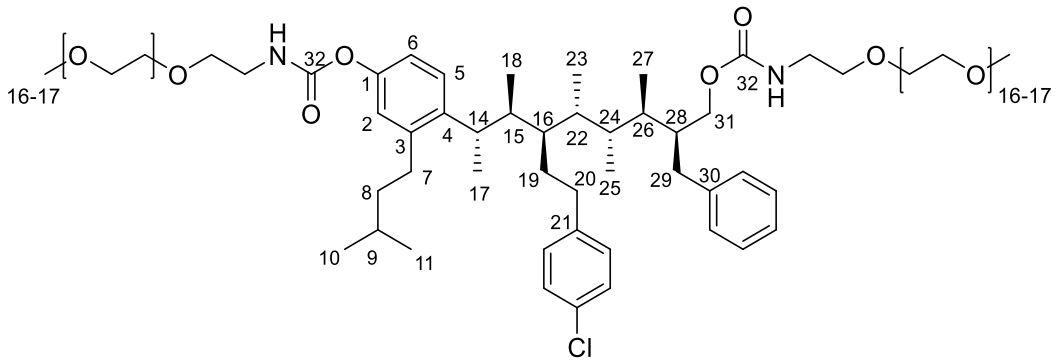


Figure 149. The ^{13}C NMR spectra of **163**

4-((12S,13R,14R,15R,16R,17S,18S)-12-benzyl-16-(4-chlorophenethyl)-13,14,15,17-tetramethyl-9-oxo-2,5,10-trioxa-8-azanonadecan-18-yl)-3-isopentylphenyl (2-(2-polymethoxyethoxy)ethyl)carbamate, 164



*Compound **164** was synthesised by Dr. Madhavachary Rudrakshula using the same method as described for **157**.*

R_f: 0.12 (1:9 MeOH:DCM)

FTIR (ν_{max}/cm^{-1}, neat): 3353, 2927, 2868, 1743, 1719, 1495, 1459, 1350, 1247, 1103, 1039

^{1}H NMR (500 MHz, MeOD) δ 7.28 (t, *J* = 7.5 Hz, 2H, Ar), 7.26-7.15 (m, 8H, Ar), 6.94 (dd, *J* = 8.5, 2.6 Hz, 1H, H_6), 6.88 (d, *J* = 2.5 Hz, 1H, H_2), 4.04-3.91 (m, 2H, H_{31}), 3.81-3.75 (m, 1H), 3.75-3.57 (m, 126H, PEG), 3.57-3.53 (m, 4H, CH_2-PEG), 3.53-3.47 (m, 2H, CH_2-PEG), 3.38 (t, *J* = 5.5 Hz, 2H, CH_2-PEG), 3.37 (s, 6H, CH_3-PEG), 3.26 (t, *J* = 5.5 Hz, 2H), 3.01-2.91 (m, 1H, H_{14}), 2.81-2.67 (m, 4H, $H_{20,7',29''}$), 2.63-2.54 (m, 1H, $H_{7''}$), 2.42-2.32 (m, 2H, $H_{29'',28}$), 2.04-1.95 (m, 1H, H_{15}), 1.87-1.73 (m, 5H), 1.66-1.56 (m, 1H, H_9), 1.52-1.42 (m, 3H), 1.19 (d, *J* = 6.8 Hz, 3H, H_{17}), 1.01-0.91 (m, 12H), 0.84 (d, *J* = 6.1 Hz, 3H, H_{23}), 0.55 (d, *J* = 6.9 Hz, 3H, H_{18})

^{13}C NMR (101 MHz, MeOD) δ: 157.7, 156.0, 148.7, 142.3, 141.5, 141.4, 140.9, 131.3, 129.5, 128.6, 128.2, 128.2, 126.7, 125.7, 121.8 (C_2), 119.3 (C_6), 71.6, 70.2, 70.2, 70.0, 69.9, 69.9, 69.6, 69.4, 65.7 (C_{31}), 57.7, 40.8, 40.7, 40.6, 40.3, 40.0, 36.2, 36.1, 36.0, 35.9, 35.5, 31.8 (C_{29}), 31.2 (C_7), 31.0 (C_{19}), 28.1 (C_9), 21.8, 21.3, 20.5 (C_{17}), 11.7, 11.4, 11.2, 10.9 (C_{23})

MALDI MS:

Table 38. The Mw distribution table obtained from MALDI for PEG derivative **164**

		$[M + Na]^+$		
n	**Calculated Mw**	**Calculated**	**Observed**	**%Intensity**
27	1981.2	2004.2	2004.4	30
28	2025.2	2048.2	2048.5	46
29	2069.2	2092.2	2092.5	52
30	2113.3	2136.3	2136.5	73
31	2157.3	2180.3	2180.6	86
32	2201.3	2224.3	2224.6	96
33	**2245.3**	**2268.3**	**2268.6**	**100**
34	2289.4	2312.4	2312.7	86
35	2333.4	2356.4	2356.7	89
36	2377.4	2400.4	2400.7	78
37	2421.4	2444.4	2444.7	50
38	2465.5	2488.5	2488.8	58
39	2509.5	2532.5	2532.8	46
40	2553.5	2576.5	2576.8	34
41	2597.5	2620.5	2620.8	30
43	2685.6	2708.6	2708.9	19

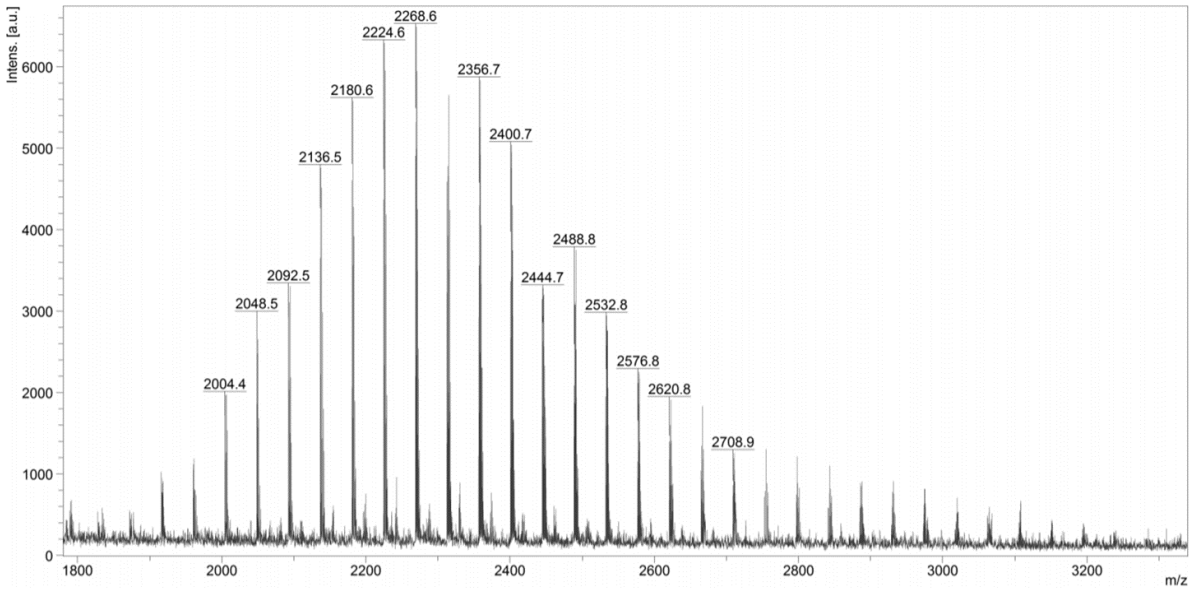


Figure 150. The MALDI MS spectrum of **164**

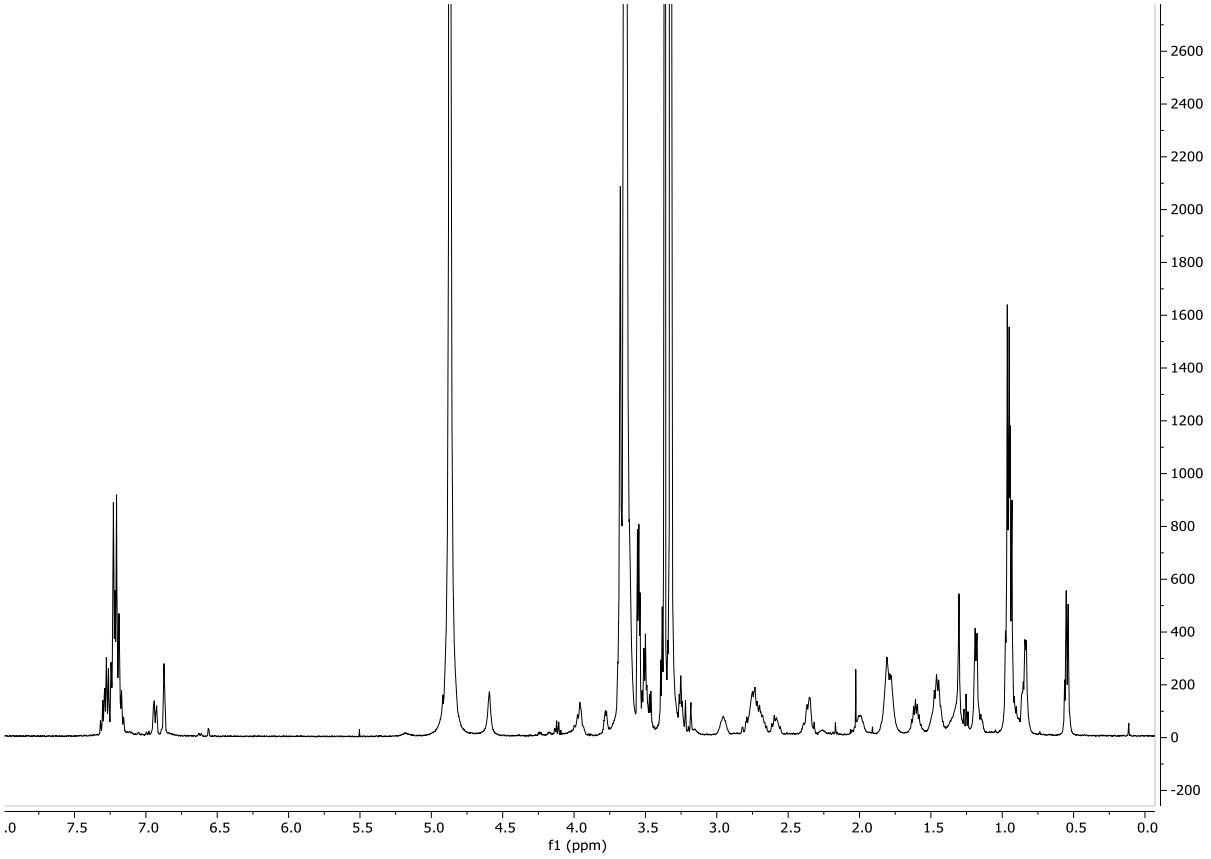


Figure 151. The ^{1}H NMR spectra of **164**

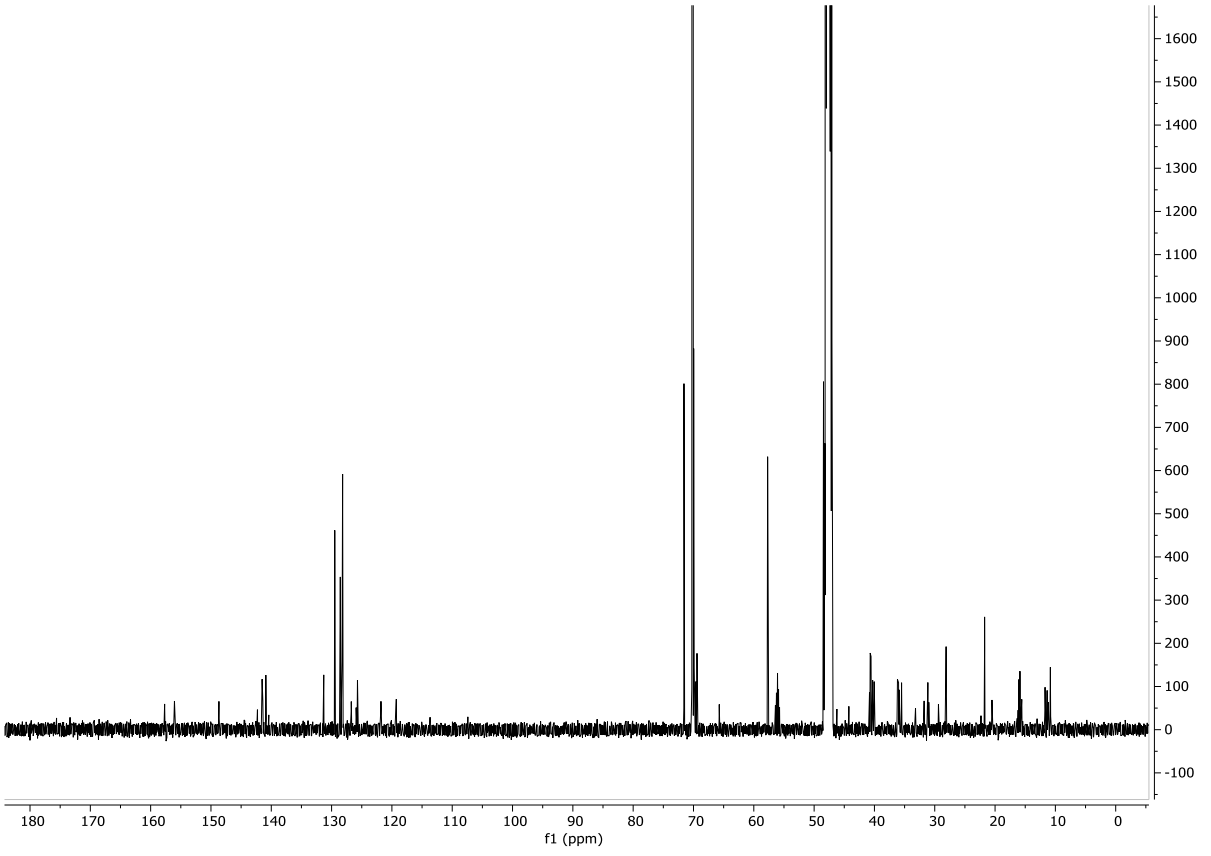
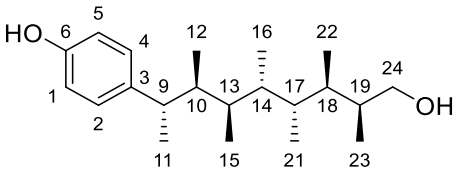


Figure 152. The ^{13}C NMR spectra of **164**

4-((2S,3S,4R,5S,6R,7R,8S)-9-hydroxy-3,4,5,6,7,8-hexamethylnonan-2-yl)phenol, 171



Compound **171** was synthesised by Dr. Madhavachary Rudrakshula using a lithiation–borylation sequence resembling **155**.

R_f: 0.22 (3:7 EtOAc:hexane)

$[\alpha]^{20}_D$: +35.0 (c 1.0, $CHCl_3$)

FTIR (ν_{max}/cm^{-1}, neat): 3438, 3119, 2957, 2920, 1594, 1514, 1381, 1254, 1225, 1059, 830

^{1}H NMR (500 MHz, MeOD) δ: 6.98 (d, J = 8.5 Hz, 2H, $H_{2,4}$), 6.71 (d, J = 8.5 Hz, 2H, $H_{1,5}$), 3.50-3.39 (m, 2H, H_{24}), 3.37 (s, 1H), 2.45 (dq, J = 10.3, 6.9 Hz, 1H, H_9), 1.97 (hd, J = 7.0, 2.2 Hz, 1H, H_{19}), 1.85-1.79 (m, 1H, H_{10}), 1.79-1.73 (m, 1H), 1.68-1.61 (m, 2H), 1.61-1.54 (m, 1H, H_{13}), 1.18 (d, J = 6.9 Hz, 3H, H_{11}), 0.82 (d, J = 5.8 Hz, 3H), 0.80 (d, J = 6.6 Hz, 3H), 0.79 (d, J = 6.6 Hz, 3H, H_{15}), 0.77 (d, J = 7.0 Hz, 3H, H_{23}), 0.73 (d, J = 6.3 Hz, 3H), 0.50 (d, J = 6.8 Hz, 3H)

^{13}C NMR (126 MHz, MeOD) δ: 154.9 (C_6), 139.0 (C_3), 127.9 ($C_{2,4}$), 114.6 ($C_{1,5}$), 66.3 (C_{24}), 43.2 (C_9), 39.8 (C_{10}), 35.8 (C_{19}), 35.2, 35.2, 35.1, 35.1, 19.6 (C_{11}), 10.8 (C_{12}), 10.2, 10.1, 10.1, 9.9, 8.8 (C_{23})

HRMS: (MALDI) calcd. for $C_{21}H_{36}O_2Na$ ($M+Na^+$): 343.2608; Found 343.2615.

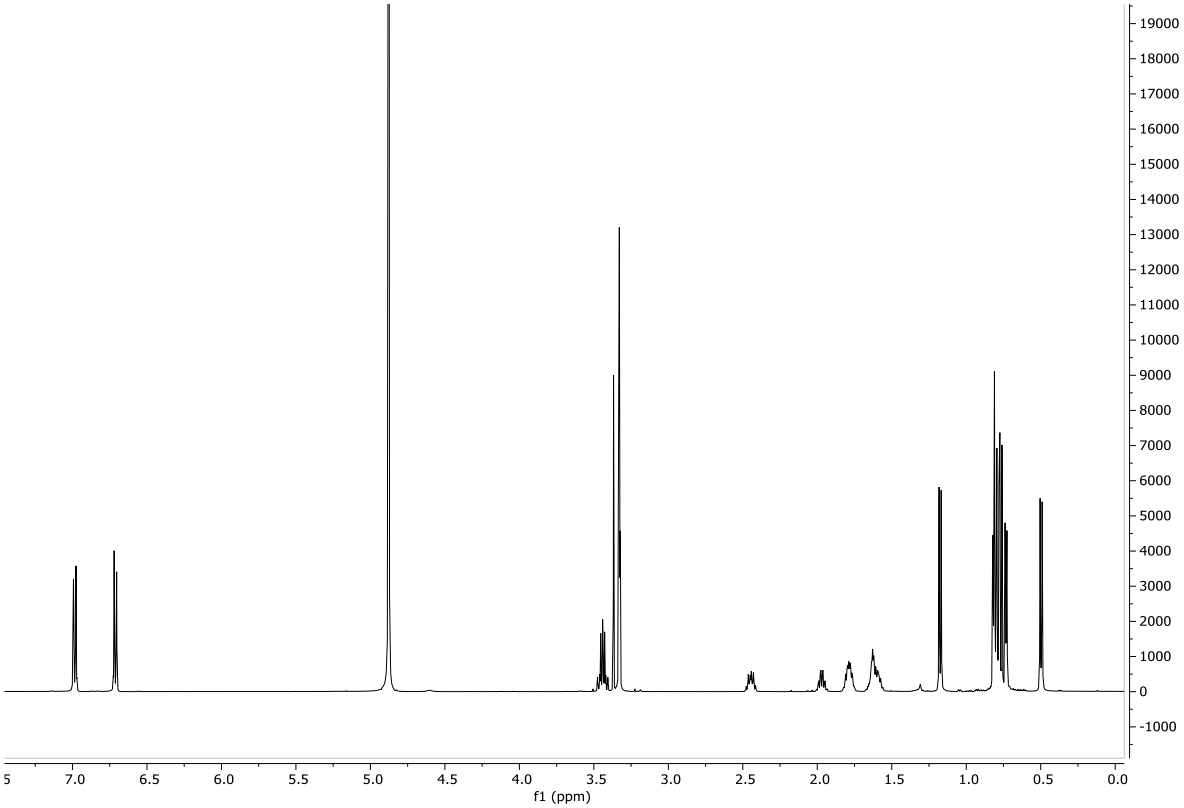


Figure 153. The ^{1}H NMR spectra of **171**

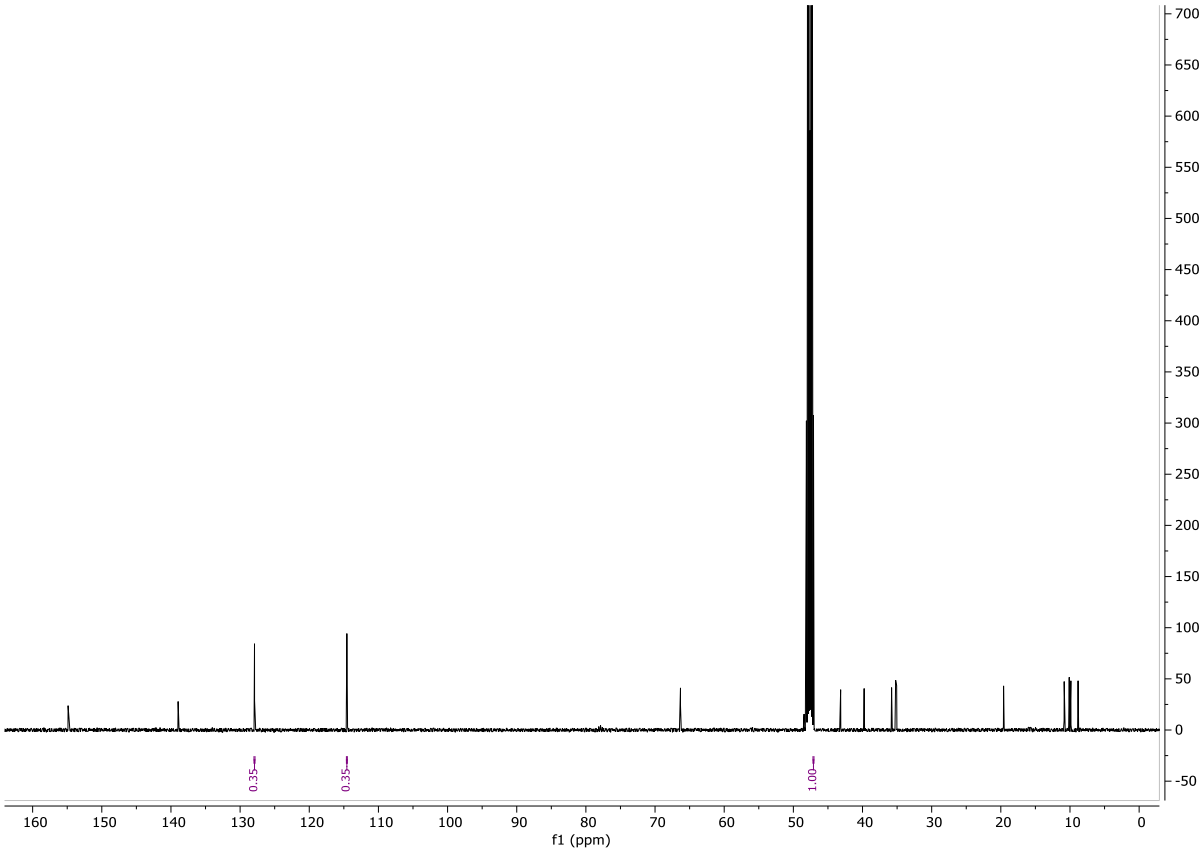
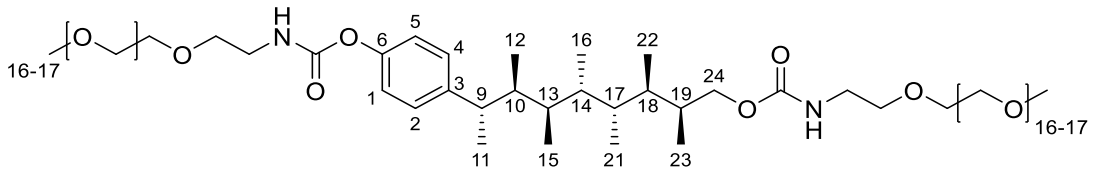


Figure 154. The ^{13}C NMR spectra of **171**

(2S,3R,4R,5S,6R,7S,8S)-8-(4-(((2-(2-methoxyethoxy)ethyl)carbamoyl)oxy)phenyl)-2,3,4,5,6,7-hexamethylnonyl poly-(2-(2-methoxyethoxy)ethyl)carbamate, 172



Compound **172** was synthesised by Dr. Madhavachary Rudrakshula using the same method as described for **157**.

R_f: 0.20 (1:9 MeOH:DCM)

FTIR (ν_{max}/cm^{-1}, neat): 3360, 2924, 2870, 1743, 1605, 1460, 1350, 1250, 1105

^{1}H NMR (500 MHz, MeOD) δ: 7.20 (d, J = 8.7 Hz, 2H, $H_{2,4}$), 7.04 (d, J = 8.7 Hz, 2H, $H_{1,5}$), 3.95 (d, J = 7.3 Hz, 2H, H_{24}), 3.84-3.76 (m, 2H, CH_2-PEG), 3.74-3.58 (m, 128H, PEG), 3.58-3.53 (m, 6H, CH_2-PEG), 3.53-3.46 (m, 2H, CH_2-PEG), 3.39 (t, J = 5.4 Hz, 2H, CH_2-PEG), 3.38 (s, 6H, CH_3-PEG), 3.30 (t, J = 5.6 Hz, 2H, H_{26}), 2.67-2.48 (m, 1H, H_9), 2.27-2.06 (m, 1H, H_{18}), 1.95-1.75 (m, 2H, $H_{10,13}$), 1.71-1.49 (m, 3H), 1.22 (d, J = 6.9 Hz, 3H, H_{11}), 0.90-0.72 (m, 15H), 0.52 (d, J = 6.8 Hz, 3H, H_{12})

^{13}C NMR (101 MHz, MeOD) δ: 157.8 (C_{25}), 156.0 (C_{28}), 149.3 (C_6), 145.1 (C_3), 127.9 ($C_{2,4}$), 121.2 ($C_{1,5}$), 71.6, 70.3, 70.2, 70.2, 70.2, 70.1, 70.0, 69.9, 69.9, 69.6, 69.4, 68. (C_{24}), 57.7, 43.5 (C_9), 40.6, 40., 39.7 (C_{10}), 35.5, 35.2, 35.1, 33.0 (C_{19}), 19.6 (C_{11}), 11.0 (C_{12}), 10.3, 10.2, 9.9, 8.9 (C_{23})

MALDI MS:

Table 39. The Mw distribution table obtained from MALDI for PEG derivative **172**

n	Calculated Mw	$[M + Na]^+$		%Intensity
		Calculated	Observed	
23	1535.0	1557.9	1558.2	20
24	1579.0	1602.0	1602.2	30
25	1623.0	1646.0	1646.3	46
26	1667.0	1690.0	1690.3	53
27	1711.1	1734.1	1734.3	63
28	1755.1	1778.1	1778.4	56
29	1799.1	1822.1	1822.4	66
30	1843.1	1866.1	1866.4	80
31	1887.1	1910.2	1910.4	86
32	**1931.2**	**1954.2**	**1954.5**	**100**
33	1975.2	1998.2	1998.5	93
34	2019.2	2042.2	2042.5	70
35	2063.3	2086.3	2086.6	53
36	2107.3	2130.3	2130.6	36
37	2195.4	2218.3	2218.6	30
39	2239.4	2262.4	2262.7	23

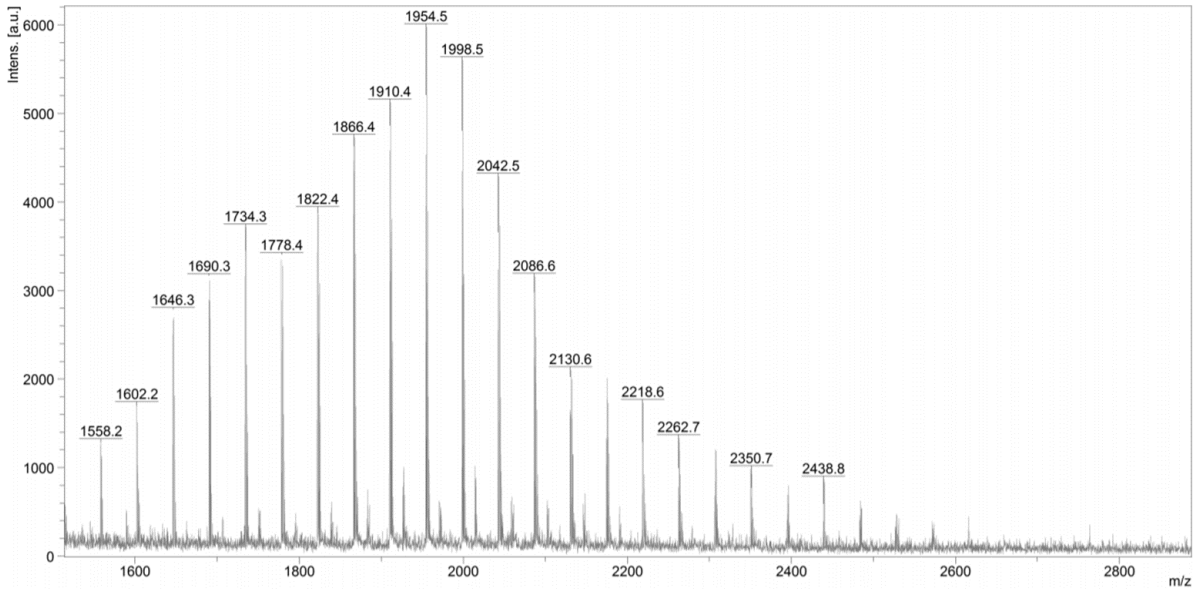


Figure 155. The MALDI MS spectrum of **172**

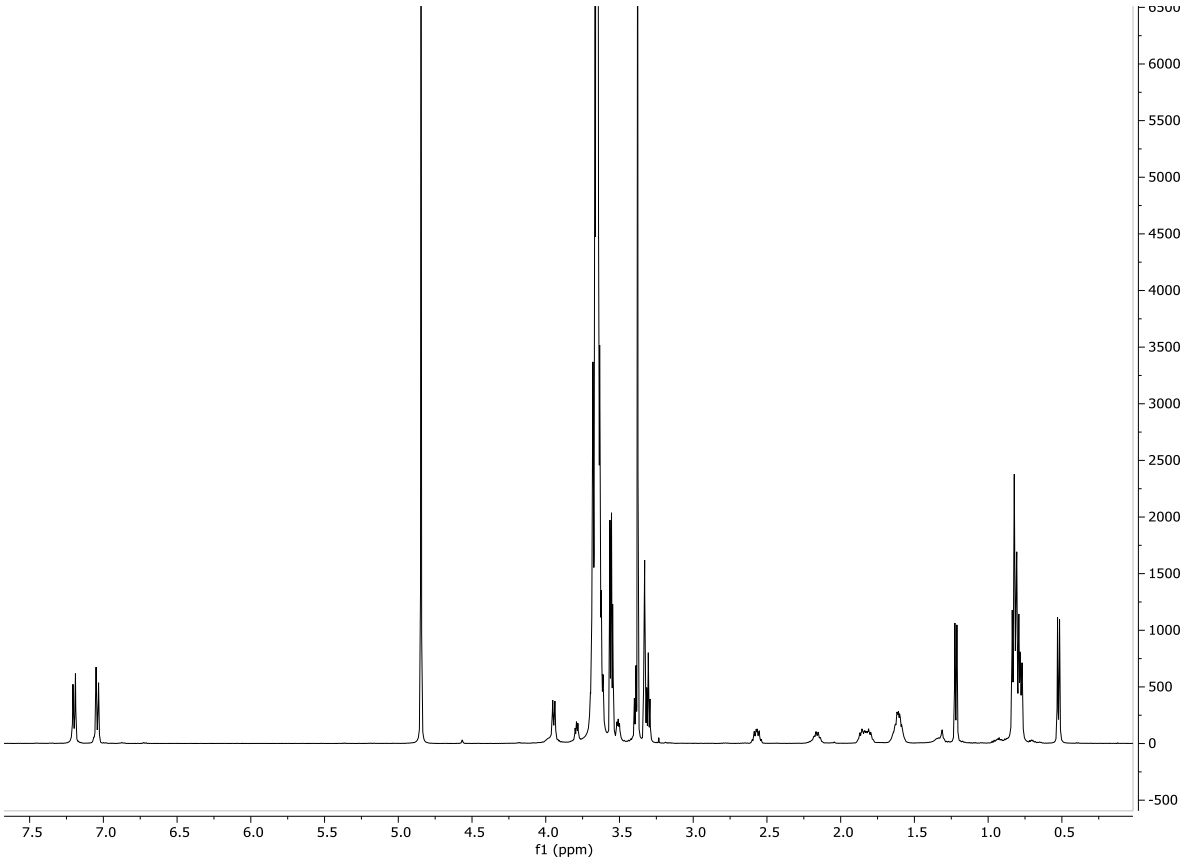


Figure 156. The ^{1}H NMR spectra of **172**

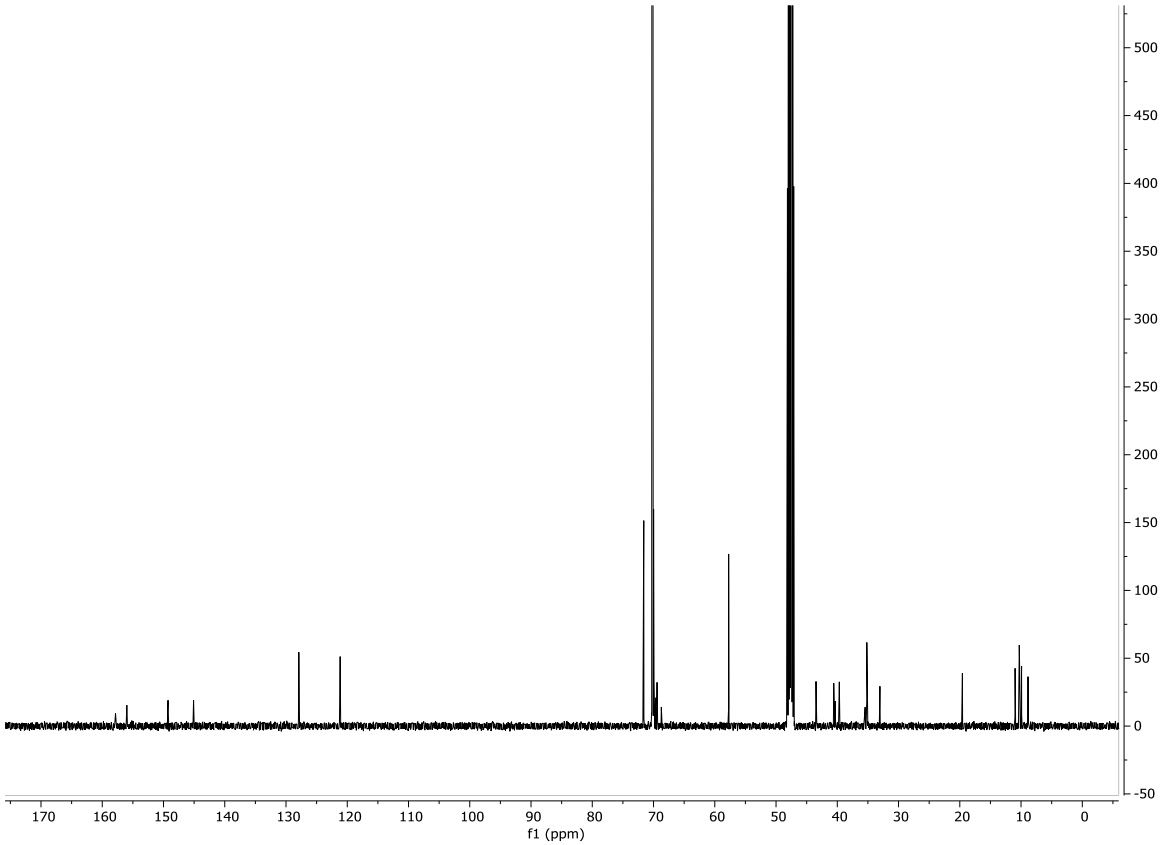
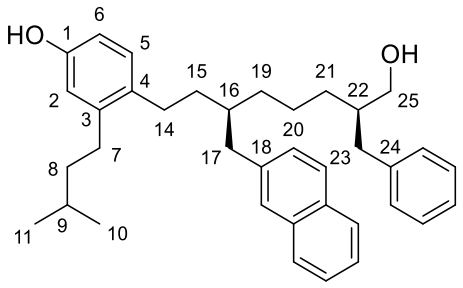


Figure 157. The ^{13}C NMR spectra of **172**

4-((3R,7R)-7-benzyl-8-hydroxy-3-(naphthalen-2-ylmethyl)octyl)-3-isopentylphenol, 178



Compound **178** was synthesised by Dr. Madhavachary Rudrakshula using a lithiation–borylation sequence resembling **155**

R_f: 0.4 (4:6 EtOAc:Hexane)

$[\alpha]^{20}_D$: –0.41 (c 1.0, MeOH)

IR (ν_{max}/cm^{-1}, neat): 3438, 3119, 2967, 2920, 1594, 1514, 1460, 1382, 1254, 1226, 1060, 830

^{1}H NMR (500 MHz, MeOD) δ: 7.81 (dd, J = 7.7, 1.2 Hz, 1H, Ar), 7.79-7.75 (m, 2H, Ar), 7.58 (s, 1H, Ar), 7.48-7.39 (m, 2H, Ar), 7.29 (dd, J = 8.3, 1.8 Hz, 1H, Ar), 7.25-7.19 (m, 2H, Ar), 7.17-7.11 (m, 3H, Ar), 6.84 (d, J = 8.1 Hz, 1H, H_5), 6.53 (d, J = 2.5 Hz, 1H, H_2), 6.50 (dd, J = 8.1, 2.5 Hz, 1H, H_6), 3.45-3.42 (m, 2H, H_{25}), 2.79 (dd, J = 13.6, 6.8 Hz, 1H, $H_{17'}$), 2.72 (dd, J = 13.6, 3.0 Hz, 1H, $H_{17''}$), 2.68 (dd, J = 9.3, 2.8 Hz, 1H, $H_{23'}$), 2.61-2.54 (m, 1H, $H_{14'}$), 2.53-2.48 (m, 1H, $H_{23''}$), 2.49-2.43 (m, 1H, $H_{14''}$), 2.39 (td, J = 7.4, 3.1 Hz, 2H, H_7), 1.87-1.79 (m, 1H, H_{16}), 1.79-1.70 (m, 1H, H_{22}), 1.56-1.18 (br m, 12H), 0.86 (d, J = 6.7 Hz, 3H), 0.84 (d, J = 6.7 Hz, 3H)

^{13}C NMR (126 MHz, MeOD) δ: 154.8 (C_1), 141.5, 140.9, 138.9, 133.7, 132.2, 131.1, 129.9 (C_5), 128.8, 127.8, 127.4, 127.3, 127.2, 127.0, 126.9, 125.4, 125.4, 124.7, 115.3 (C_2), 112.3 (C_6), 63.7 (C_{25}), 42.4 (C_{22}), 40.5, 40.5, 39.2 (C_{16}), 37.1 (C_{23}), 35.3, 33.4, 30.4, 30.3, 29.0 (C_{14}), 27.9 (C_9), 23.5, 21.6

HRMS: (MALDI) calcd. for $C_{37}H_{46}O_2Na$ ($M+Na^+$): 545.3990; Found 545.3999

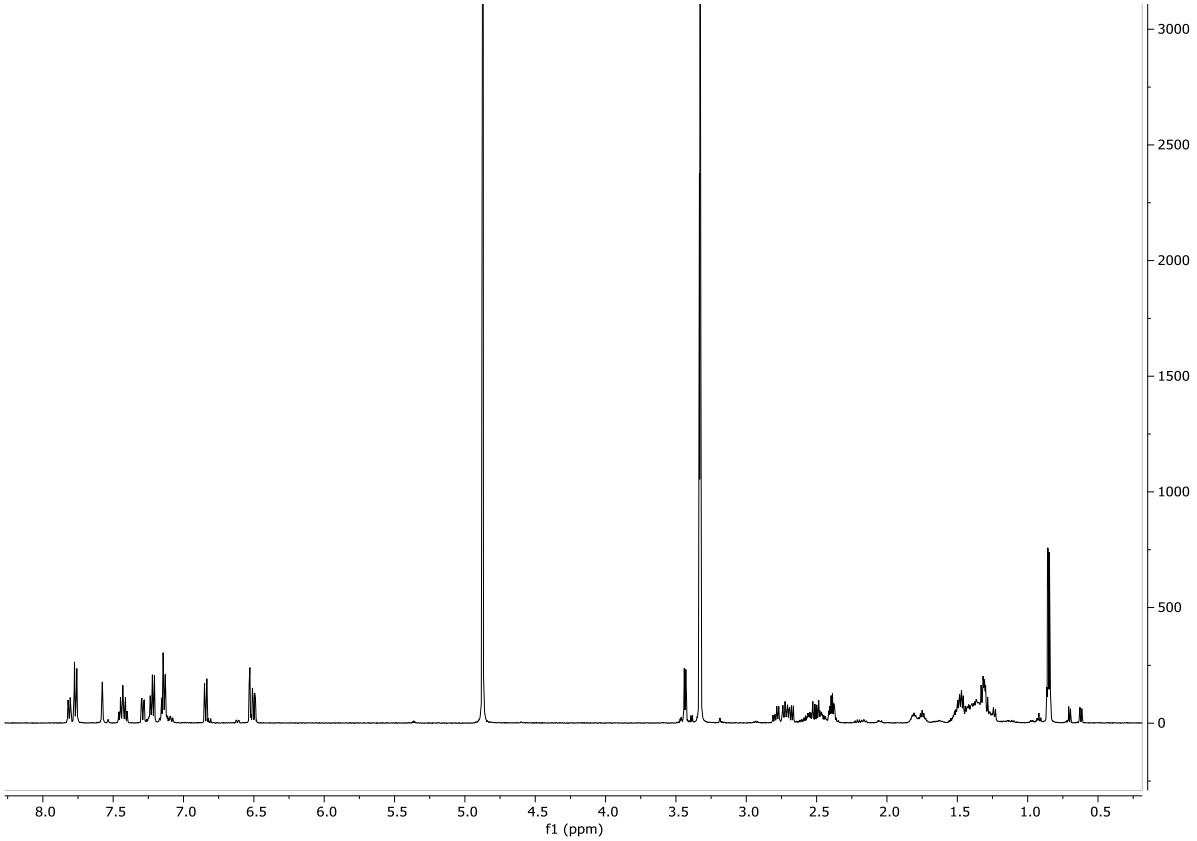


Figure 158. The ^{1}H NMR spectra of **178**

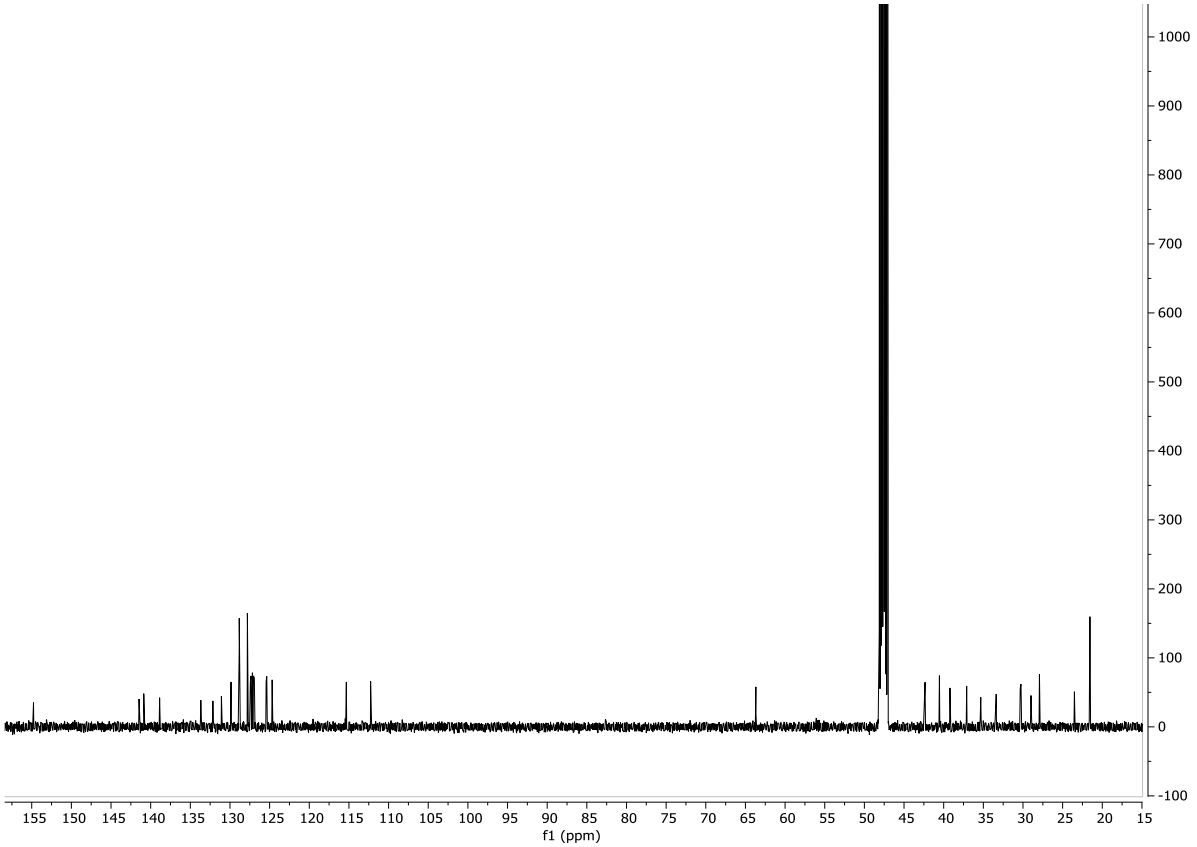
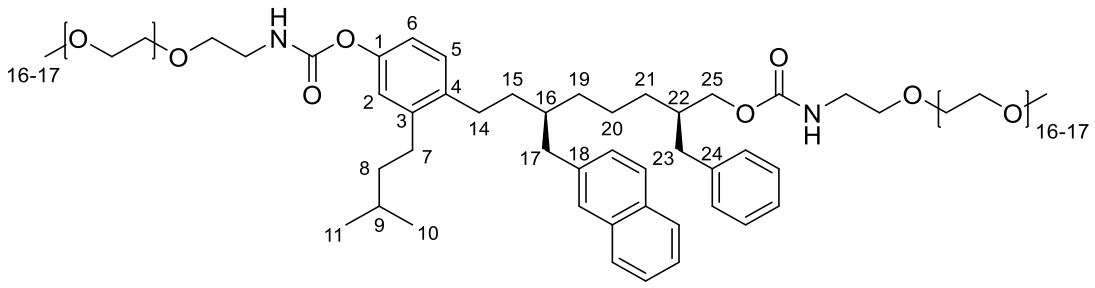


Figure 159. The ^{13}C NMR spectra of **178**

4-((3R,7R)-7-benzyl-8-hydroxy-3-(naphthalen-2-ylmethyl)octyl)-3-isopentylphenol, 179



Compound **179** was synthesised by Dr. Madhavachary Rudrakshula using the same method as described for **157**.

R_f: 0.25 (1:9 MeOH:DCM)

FTIR (ν_{max}/cm^{-1}, neat): 3431, 2921, 1610, 1514, 1462, 1382, 1225, 1060

^{1}H NMR (500 MHz, Methanol-d_4) δ: 8.56 (s, 2H, Ar), 7.83 (d, J = 7.8 Hz, 1H, Ar), 7.79 (d, J = 8.1 Hz, 2H, Ar), 7.61 (s, 1H, Ar), 7.48-7.41 (m, 2H, Ar), 7.31 (d, J = 8.4 Hz, 1H, Ar), 7.25 (t, J = 7.5 Hz, 2H, Ar), 7.20-7.12 (m, 3H, Ar), 7.04 (d, J = 8.0 Hz, 1H, H_5), 6.85-6.78 (m, 2H, $H_{1,6}$), 3.97-3.89 (m, 2H, H_{25}), 3.72-3.57 (m, 110H, PEG), 3.57-3.51 (m, 6H, CH_2-PEG), 3.37 (s, 6H, CH_3-PEG), 2.86-2.81 (m, 1H, H_{17}), 2.77-2.54 (m, 5H), 2.50-2.43 (m, 2H, H_7), 2.08-2.02 (m, 1H), 1.97-1.90 (m, 1H, H_{22}), 1.89-1.83 (m, 1H, H_{16}), 1.57-1.43 (m, 4H), 0.87 (d, J = 6.7 Hz, 3H), 0.86 (d, J = 6.7 Hz, 3H)

^{13}C NMR (126 MHz, Methanol-d_4) δ: ^{13}C NMR (126 MHz, MeOD) δ 165.4, 149.2, 141.6, 140.2, 138.6, 137.2, 133.7, 132.2, 129.7, 128.8, 128.0, 127.5, 127.3, 127.0, 125.6, 124.9, 121.7, 118.7, 71.6, 70.1, 69.9, 68.8, 66.8, 57.7, 40.5, 40.3, 39.2, 30.1, 29.3, 27.9, 23.4, 21.6

MALDI MS:

Table 40. The Mw distribution table obtained from MALDI for PEG derivative **179**

		$[M + Na]^+$		
n	**Calculated Mw**	**Calculated**	**Observed**	**%Intensity**
27	1781.1	1804.1	1804.4	11
28	1825.1	1848.1	1848.4	26
29	1869.1	1892.1	1892.4	34
30	1913.1	1936.1	1936.4	46
31	1957.2	1980.2	1980.5	76
32	2001.2	2024.2	2024.5	80
33	**2045.2**	**2068.2**	**2068.5**	**100**
34	2089.2	2112.2	2112.6	97
35	2133.3	2156.3	2156.6	96
36	2177.3	2200.3	2200.6	84
37	2221.3	2244.3	2244.6	61
38	2265.4	2288.3	2288.7	42
39	2309.4	2332.4	2332.7	30
40	2353.4	2376.4	2376.7	13
41	2397.4	2420.4	2420.8	7

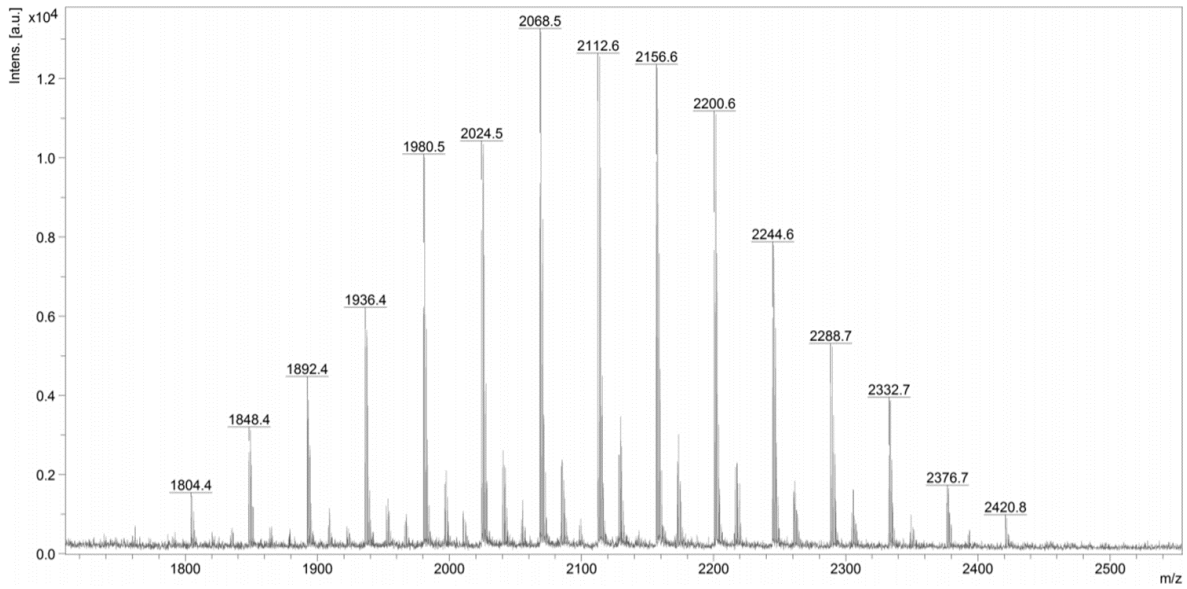


Figure 160. The MALDI MS spectrum of **179**

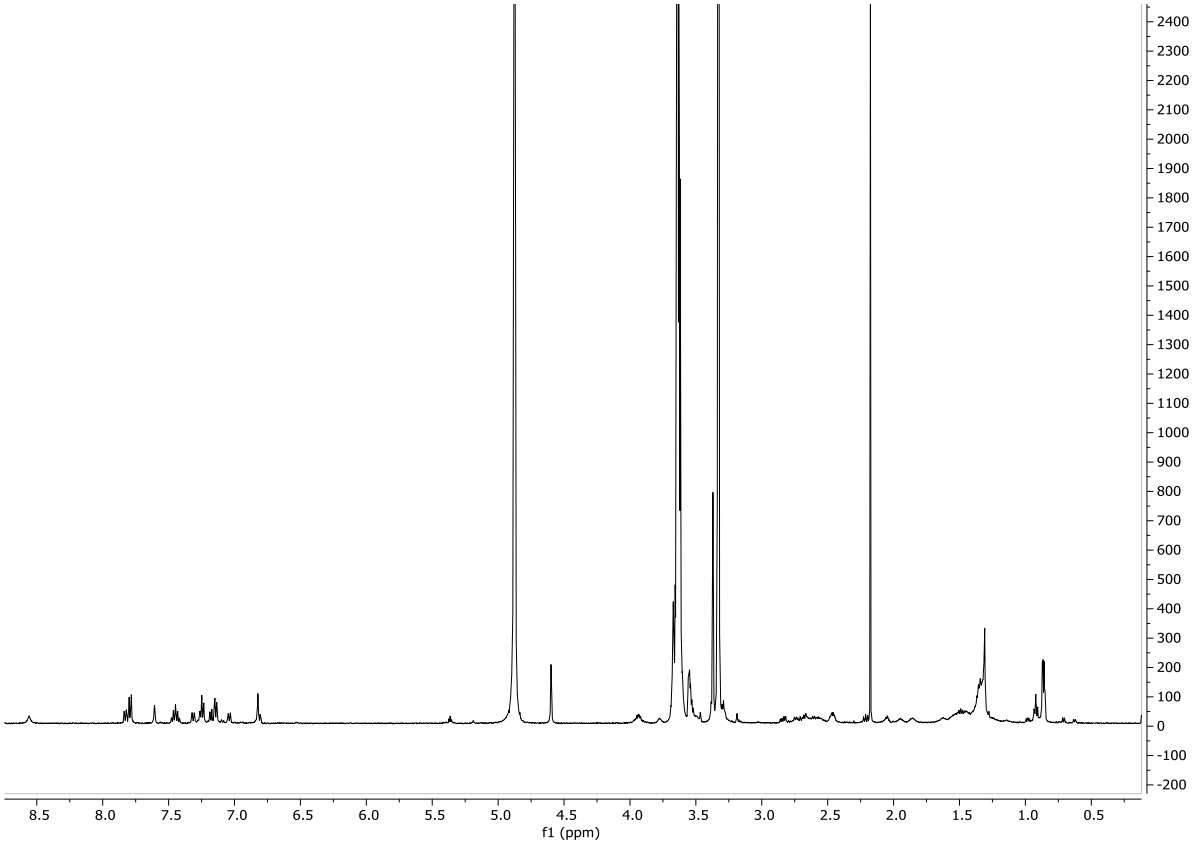


Figure 161. The ^{1}H NMR spectra of **179**

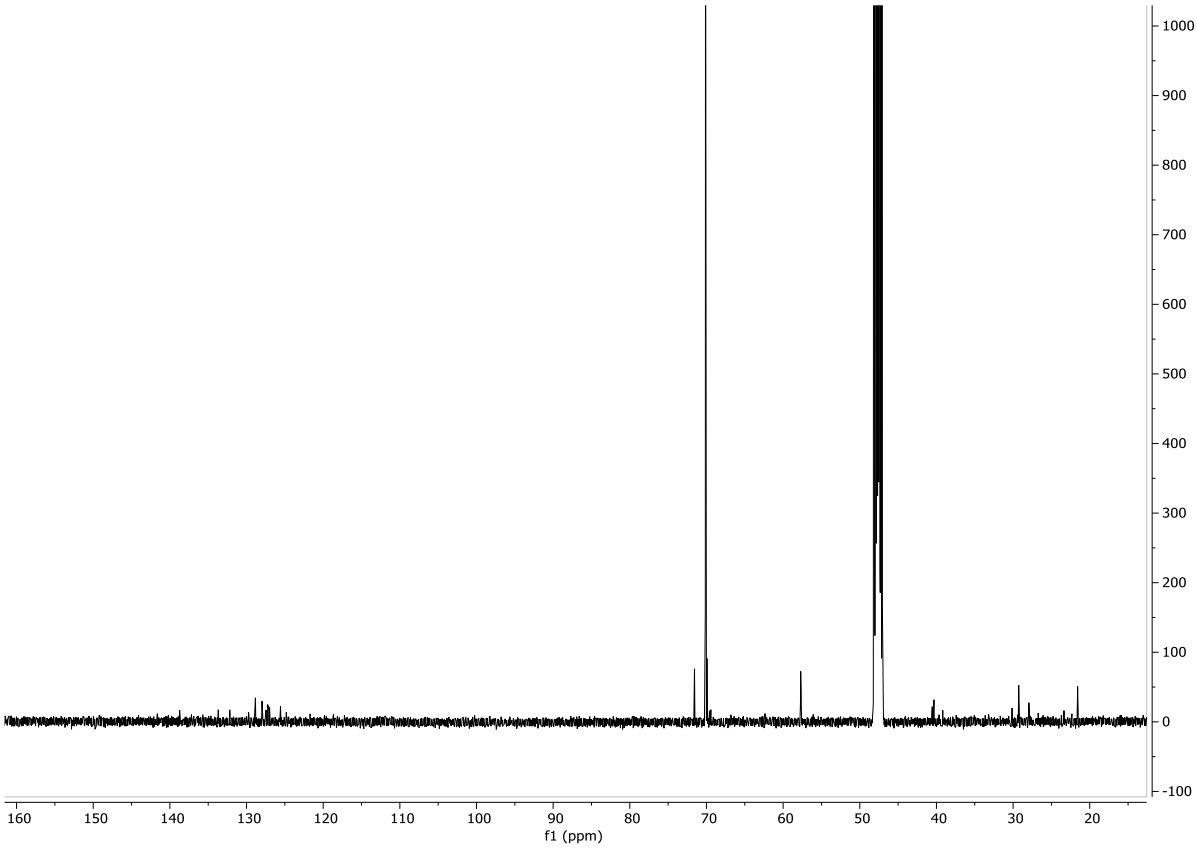


Figure 162. The ^{13}C NMR spectra of **179**

9.4 The Calculation of Gibbs Free Energies, Boltzmann Populations and Boltzmann Averaged NMR Parameters using Density Functional Theory

All raw computational data is available in \\ads.bris.ac.uk\folders\Science\Chemistry\Chemistry filestore (chm-fs)\MyShared_nmr\shareall\Computational Data_Thesis.

9.4.1 General Computational Information

DFT Calculations, including geometry optimisations, frequency calculations and NMR calculations, were performed using the Gaussian 09 Software Package.[326] DFT calculations were performed using BlueCrystal Phase 3, a supercomputing facility available within the advanced computing research center at the University of Bristol (http://www.bris.ac.uk/acrc).

Following the conformational search, the top 99.9% of conformers, corresponding to 241 conformers for **119**, were subjected to a DFT geometry optimisation-frequency calculation using mPW1PW91/6-31g(d) basis set (basis set 1-BS1). The integral equation formalism polarisable continuum model was used for solvation (IEFPCM, chloroform). Conformers that either did not converge or converged to an imaginary frequency were eliminated from the conformational pool.

The optimised geometries were subjected to a frequency single point calculation using mPW1PW91/6-311g(d,p) (basis set 2-BS2) to obtain a more accurate description of conformer energies. The calculations were performed using the same solvation method in chloroform. The Gibbs free energies (in kJ mol^{-1}) of each conformer were estimated according to **Equation 14**:

$$G_i^{BS2} = E_i^{BS2} + \left(G_i^{BS1} - E_i^{BS1}\right)$$

Equation 14. The estimation of the Gibbs energy of conformer i.

Where G_i^{BS2} is the estimated Gibbs energy of conformer i using 6-311g(d,p), E_i^{BS2} is the potential energy of conformer i calculated using 6-311g(d,p), G_i^{BS1} is the calculated Gibbs energy of conformer i using 6-31g(d) and E_i^{BS1} is the potential energy of conformer i calculated using 6-31g(d).

The estimated Gibbs energies of each conformer were subsequently used in the Boltzmann analysis of conformer population according to **Equation 15**:

$$P_i = 100 \text{ x } \left(\frac{ge^{\frac{\Delta G_i^{BS2}}{RT}}}{\sum_{i=1}^{n} ge^{\frac{\Delta G_i^{BS2}}{RT}}}\right)$$

Equation 15. The calculation of the Boltzmann population of conformer i

Where P_i is the Boltzmann population of conformer i, ΔG_i^{BS2} is the relative estimated Gibbs energy in kJ mol^{-1} of conformer i, R is the ideal gas constant (8.314 x 10^{-3} kJ mol^{-1}), T is the temperature of the system (298 K) and g is the degeneracy of conformer i (which is 1 for all conformers in this study).

The Boltzmann distribution was used to select the top 99% of conformers to subject to the DFT NMR calculations to calculate scalar coupling constants and magnetic shielding tensors. The NMR properties were computed using the GIAO (gauge-independent atomic orbitals) method using mPW1PW91/6-311g(d,p). The IEFPCM solvation model was used with the solvent chloroform chosen. The additional keyword 'mixed' was included to allow for mixing between the core orbitals and the valence orbitals.

9.4.2 The Calculation of Boltzmann Averaged Chemical Shifts

The Boltzmann averaged magnetic shielding tensors were calculated according to **Equation 16**.

$$\sigma^x = \frac{\sum_i (\sigma_i^x \text{ x } p_i)}{\sum p}$$

Equation 16. The calculation of the Boltzmann averaged magnetic shielding tensors.

Where σ^x is the magnetic shielding tensor of nucleus x for conformer i, p_i is the Boltzmann population of conformer i and $\sum p$ is the combined population of conformers involved.

The chemical shifts were calculated according to **Equation 17**.

$$\delta_{calc}^x = \frac{\sigma^{TMS} - \sigma^x}{1 - \frac{\sigma^{TMS}}{10^6}}$$

Equation 17. The calculation of chemical shift from magnetic shielding tensor

Where δ_{calc}^x is the calculated chemical shift of nucleus x, σ^x is the magnetic shielding tensor of nucleus x, σ^{TMS} is the magnetic shielding tensor of the relevant nuclei of TMS calculated in the same solvent and at the same level of theory. For proton this corresponds to 31.92258 and for carbon this corresponds to 189.3625.

The calculated chemical shifts were linearly scaled to remove any systematic errors according to **Equation 18**

$$\delta_{scaled} = (\delta_{calculated} - \text{intercept})/\text{slope}$$

Equation 18. The scaling of calculated chemical shifts

Where the intercept and the slope were obtained by plotting the calculated chemical shifts ($\delta_{calculated}$) against the experimental chemical shifts.

9.4.3 The Calculation of Boltzmann Averaged Scalar Coupling Constants

The calculated scalar coupling constants for each conformer, both HH and HC, were Boltzmann averaged against their calculated Boltzmann populations according to **Equation 19**.

$$^{n}J_{HX,calc} = \frac{\sum_i (J_{HX,i} \times p_i)}{\sum p}$$

Equation 19. The calculation of the Boltzmann averaged scalar coupling constants

Where $^{n}J_{HX,calc}$ is the Boltzmann averaged scalar coupling constant in Hz, $J_{HX,i}$ is the scalar coupling constant of conformer i, p_i is the population of conformer i and $\sum p$ is the combined population of conformers involved.

9.4.4 The Calculation of Boltzmann Averaged Interproton Distances

The interproton distances were obtained from the optimised geometries after the geometry optimisation and frequency calculation, performed using mPW1PW91/6-31g(d) basis set. The interproton distances of each conformer were Boltzmann averaged against their calculated Boltzmann populations according to **Equation 20**.

$$r_{Ha-Hb,calc} = \left(\sum_i (r_{Ha-Hb,i})^{-6} \times p_i \right)^{-\frac{1}{6}}$$

Equation 20. The calculation of Boltzmann averaged interproton distances

Where $r_{Ha-Hb,calc}$ is the Boltzmann averaged interproton distance between H_a and H_b, $r_{Ha-Hb,i}$ is the calculated interproton distance between H_a and H_b for conformer i and p_i is the population of conformer i.

9.4.5 The DFT Calculated Gibbs free energies and Boltzmann Populations for Conformers of Structure 119

The method described above was used to calculate the Gibbs energies of the conformers of **119** and their corresponding Boltzmann populations. The relative potential and Gibbs energies calculated using both basis set 1 and basis set 2, for all conformers under 10 kJ mol^{-1} are provided in **Table 41**. The relative Gibbs energies estimated for basis set 2 were used to calculate the conformer Boltzmann populations.

Table 41. The Calculated Data for **119**

Conformer	ΔE (MM) (kJ mol^{-1})	ΔE (BS1) (kJ mol^{-1})*	ΔG (BS1) (kJ mol^{-1})*	ΔE (BS2) (kJ mol^{-1})**	ΔG_{est} (BS2) (kJ mol^{-1})**	Boltzmann Population (%)
2	0.33	2.38	0.00	1.66	0.00	23.99
3	1.57	5.06	0.85	3.73	0.24	21.74
1	0.00	3.18	2.93	1.86	2.34	9.34
4	3.95	3.12	3.62	1.62	2.84	7.62
12	7.08	7.17	5.03	5.84	4.43	4.02
189	18.73	2.64	4.36	2.33	4.77	3.50
14	7.27	7.80	5.47	6.38	4.78	3.49
30	10.39	7.72	5.62	6.39	5.01	3.17
9	5.66	5.32	6.42	4.07	5.90	2.21
18	8.24	5.88	7.10	4.46	6.40	1.81
41	11.82	12.77	7.25	11.19	6.40	1.81
67	13.90	3.91	6.45	3.36	6.63	1.65
8	4.93	6.28	7.32	4.95	6.71	1.60
19	8.76	0.94	5.93	1.38	7.09	1.37
11	7.02	0.00	6.86	0.00	7.59	1.12
147	17.38	6.13	7.56	5.93	8.09	0.92
110	16.24	6.85	8.08	6.33	8.28	0.85
86	14.93	7.33	8.71	7.16	9.26	0.57
188	18.72	7.37	9.30	7.25	9.91	0.44
131	16.87	12.88	9.98	12.32	10.15	0.40

*mPW1PW91/6-31g(d)

**mPW1PW91/6-311g(d,p)

The Boltzmann populations calculated by DFT can be used to plot a 3D-bubble plot of dihedral number, dihedral angle and population (**Figure 163**). The size of each bubble represents the population of the conformer contributing to that dihedral angle. The total population of conformers adopting a linear conformation after DFT is 81%.

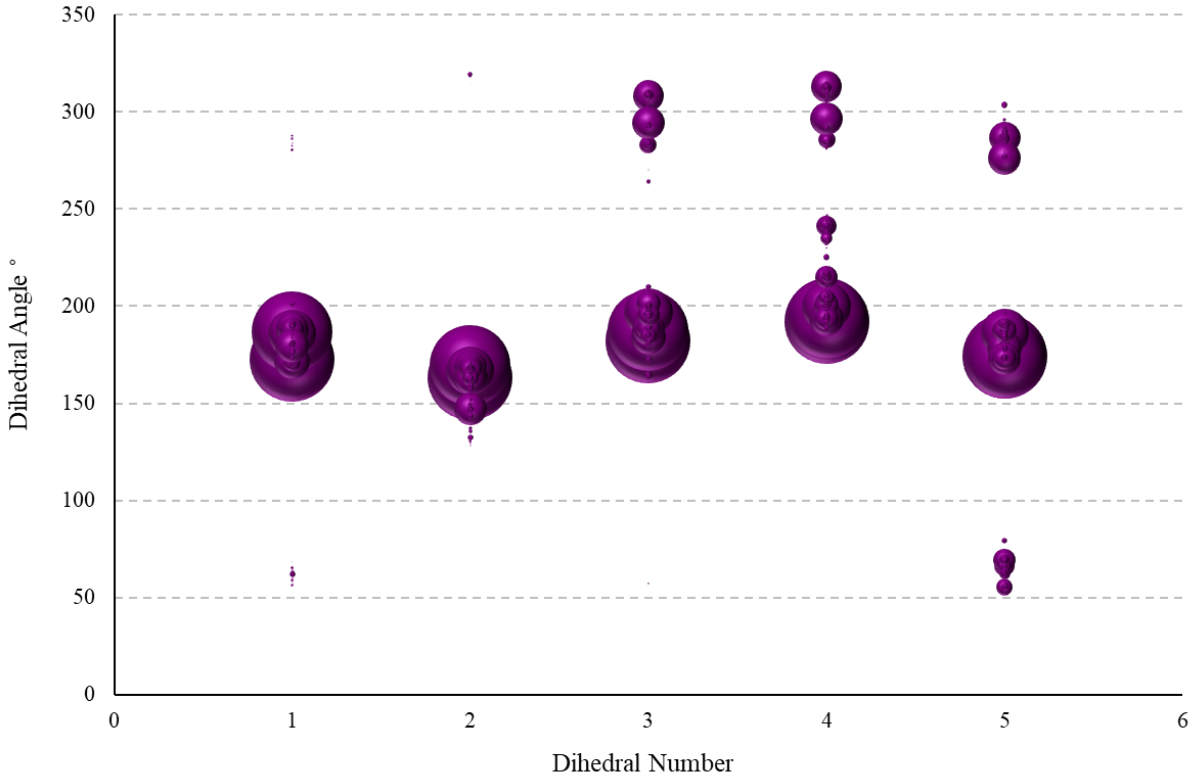


Figure 163. The bubble plot of **119** using populations calculated from DFT calculations.

The superposition of the three lowest energy conformers, which have a combined population of 55%, show very similar backbones with changes only in the side chains (**Figure 164**).

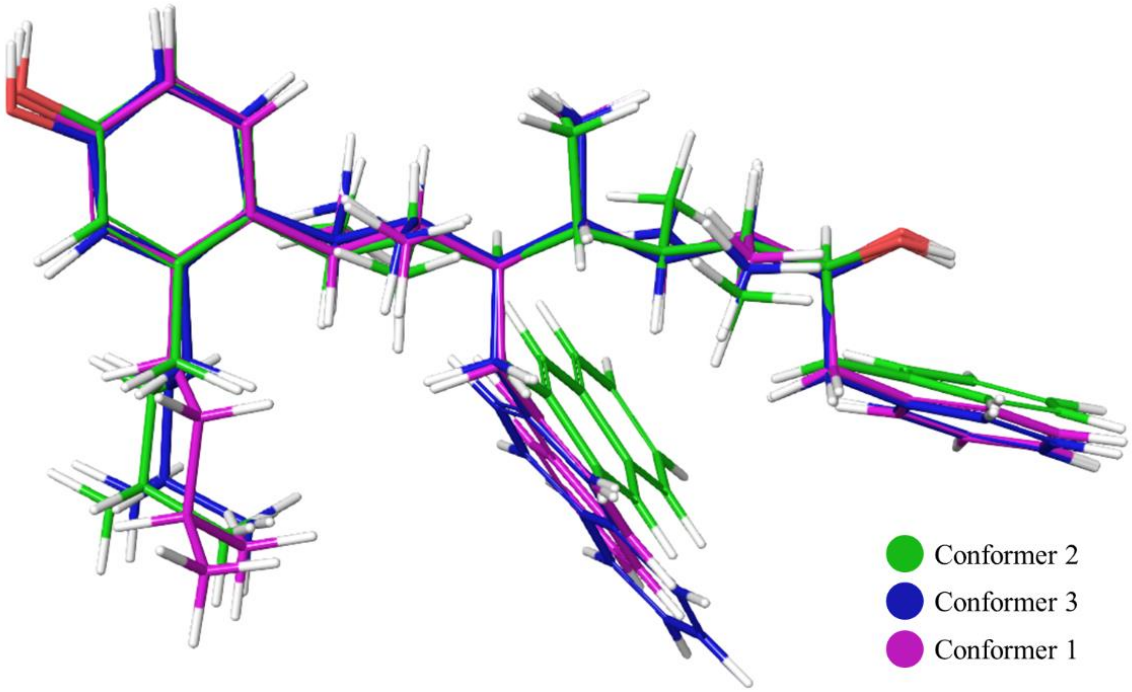


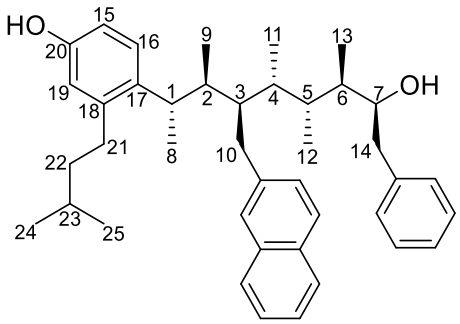
Figure 164. The three lowest energy conformers of **119**

9.5 The Acquisition of NMR Parameters in Conformational Analysis

9.5.1 The Measurement of $^{n}J_{HH}$ Scalar Coupling Constants

$^{n}J_{HH}$ scalar coupling constants were measured either directly from multiplets in the 1D ^{1}H spectrum or from simulations of the ^{1}H NMR spectrum using Spin Simulation in MestreNova. Prior to Spin Simulation, a PSYCHE pureshift ^{1}H spectrum was obtained to provide exact chemical shifts. The measured chemical shifts, estimated J-couplings and line widths of the peaks were entered into the Spin Simulation software and the spectrum was simulated (**Table 42**, **Table 43**). The simulated spectrum (**Figure 165**) was superimposed with the experimental spectrum (**Figure 166**) and the J-couplings were changed in steps of 0.1 Hz until the best fit between the experimental and simulated spectrum was observed.

Table 42. The parameters used in the partial simulation of the spectrum of **119**



Atom	Chemical Shift	N	Spin	Line Width
1	2.9048	1	½	3.8
2	2.0221	1	½	4.5
3	2.3428	1	½	4.5
4	1.8814	1	½	3.0
5	1.6855	1	½	2.8
6	1.6309	1	½	3.0
7	3.5727	1	½	2.5
8	1.2666	3	½	2.0
9	0.6955	3	½	2.0
10’	3.0223	1	½	3.0
10’	2.6628	1	½	3.5
11	0.8801	3	½	2.0
12	0.6967	3	½	2.5
13	0.9281	3	½	2.5
14’	2.4323	1	½	3.0
14’’	2.2415	1	½	2.5

Table 43. The coupling constants used in the simulation of the spectrum of **119**

	1	2	3	4	5	6	7	8	9	10’	10’’	11	12	13	14’	14’’
1																
2	10.3															
3		2.0														
4			8.2													
5				2.9												
6					9.2											
7						4.6										
8	6.7															
9		6.6														
10’			3.1													
10’’			8.4							15.1						
11				6.9												
12					6.8											
13						6.8										
14’							2.2									
14’’							9.8								13.4	

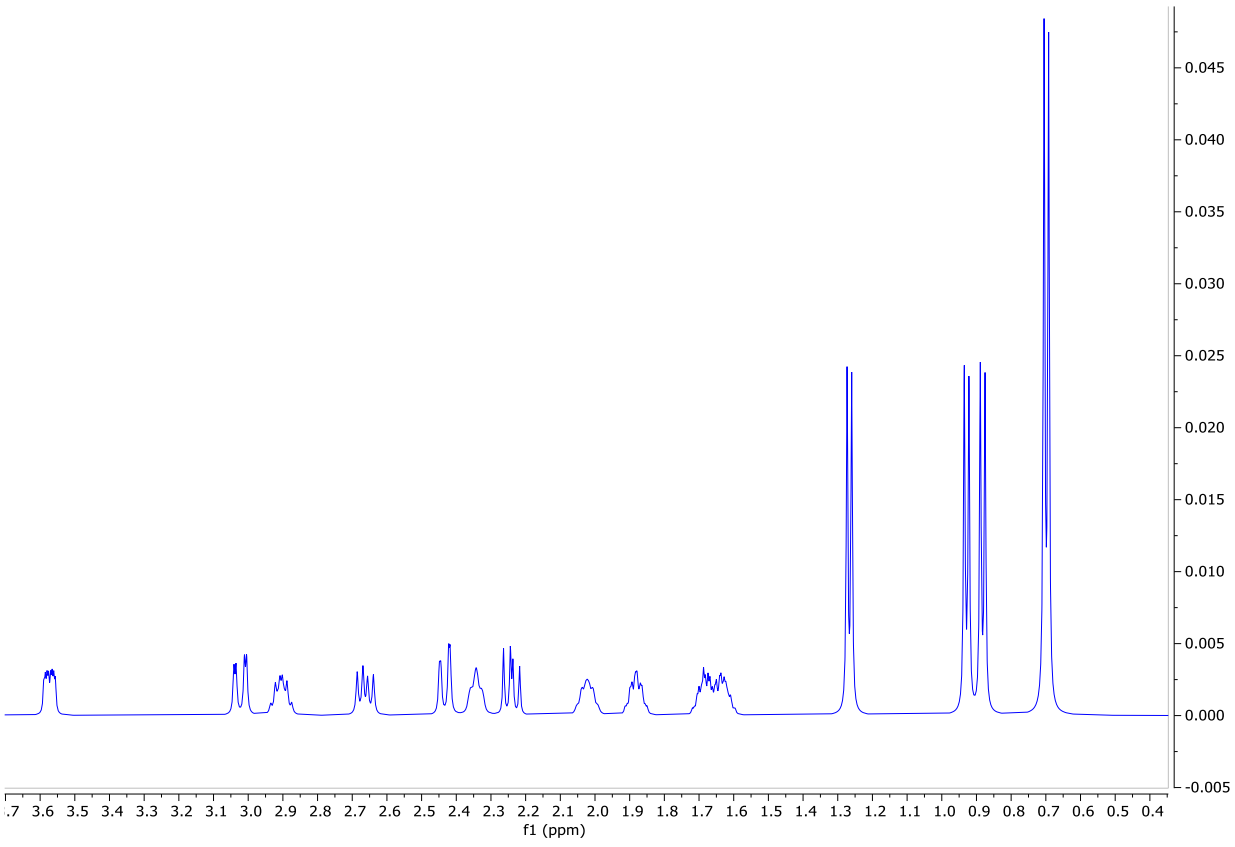


Figure 165. The partial simulation of the spectrum of **119**

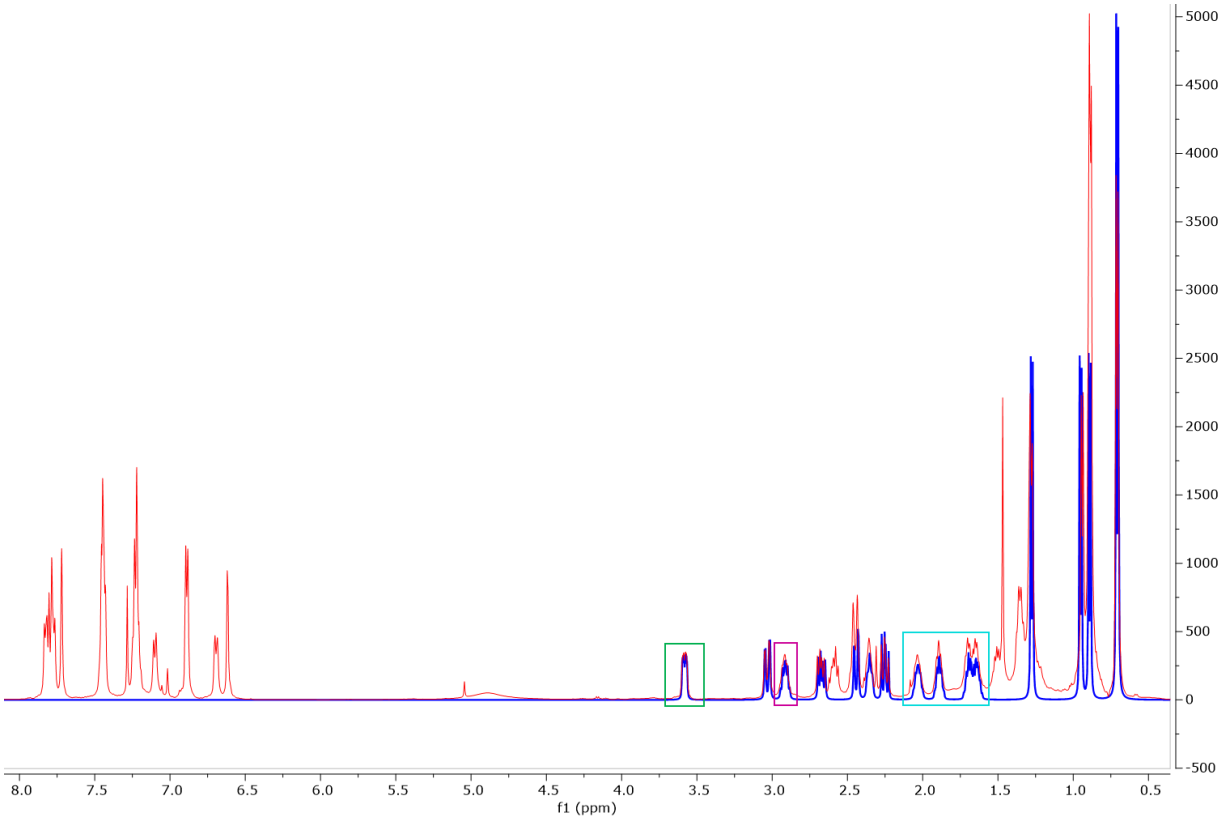


Figure 166. The overlay of the simulated spectrum (blue) and the experimental spectrum (red). Some of the backbone peaks are highlighted by coloured boxes and a zoom in of these peaks is shown in the subsequent figures.

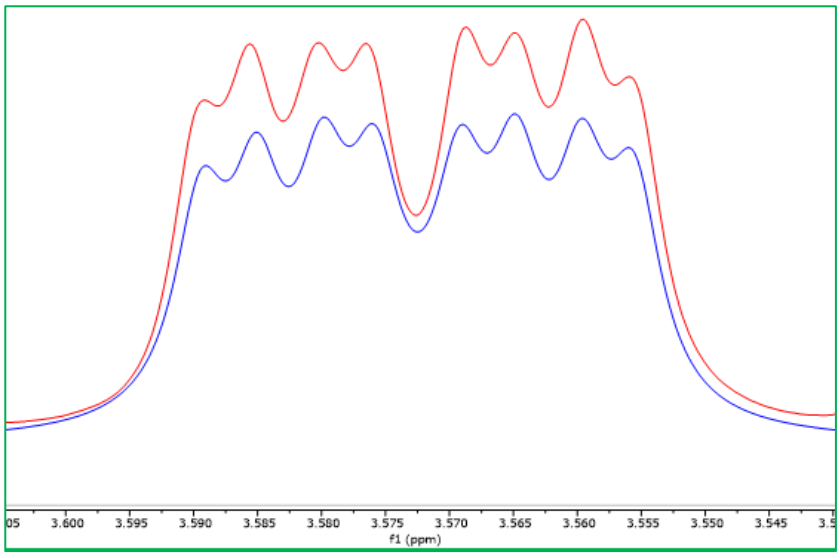


Figure 167. The simulated (blue) and experimental (red) peak of H_7

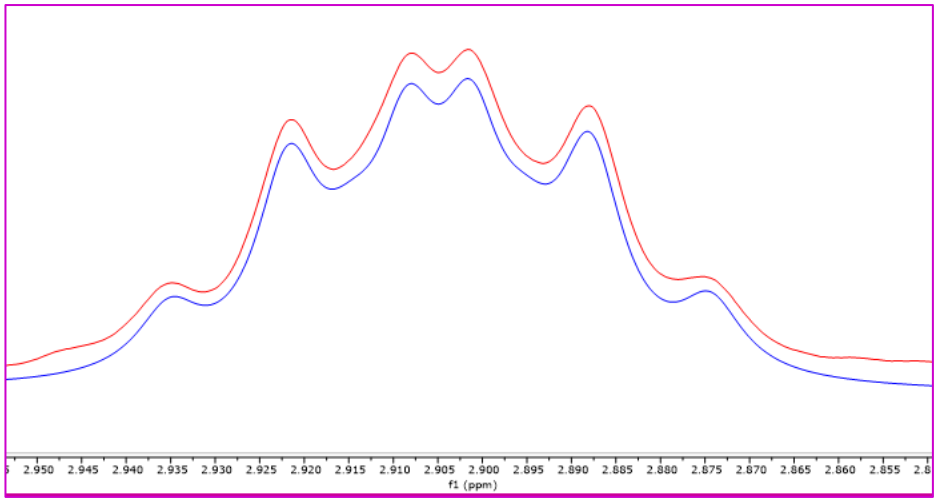


Figure 168. The simulated (blue) and experimental (red) peak of H_1.

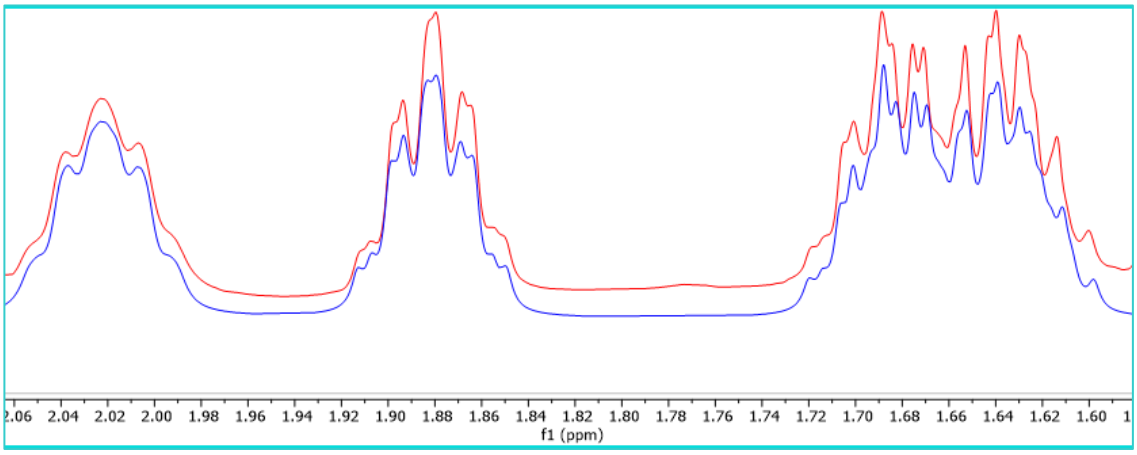


Figure 169. The simulated (blue) and experimental (red) peaks of H_2, H_4, H_5 and H_6.

9.5.2 The Measurement of $^{n}J_{CH}$ Scalar Coupling Constants

$^{n}J_{CH}$ coupling constants were measured using Accordion In-Phase and Anti Phase (IPAP) HSQMBC NMR Spectra. The In-Phase and Anti Phase spectra were acquired interleaved. The acquisition parameters include: 8 scans, 2048 f1 increments and 16384 f2 data points, f1 spectral width of 180 ppm (22624 Hz), f2 spectral width of 11 ppm (5507 Hz). The spectrum was acquired in $CDCl_3$ with 9 mg in 0.7 mL. The sum (**Figure 170**) and difference (**Figure 171**) were created by adding and subtracting the IP and AP spectra. The offset between the sum and the difference multiplets allows the extraction of the values of $^{n}J_{CH}$ (**Figure 172**). The data was measured using a Bruker AVANCE III HD 500 MHz NMR spectrometer with a 5 mm DCH ^{13}C-^{1}H/D Cryo Probe

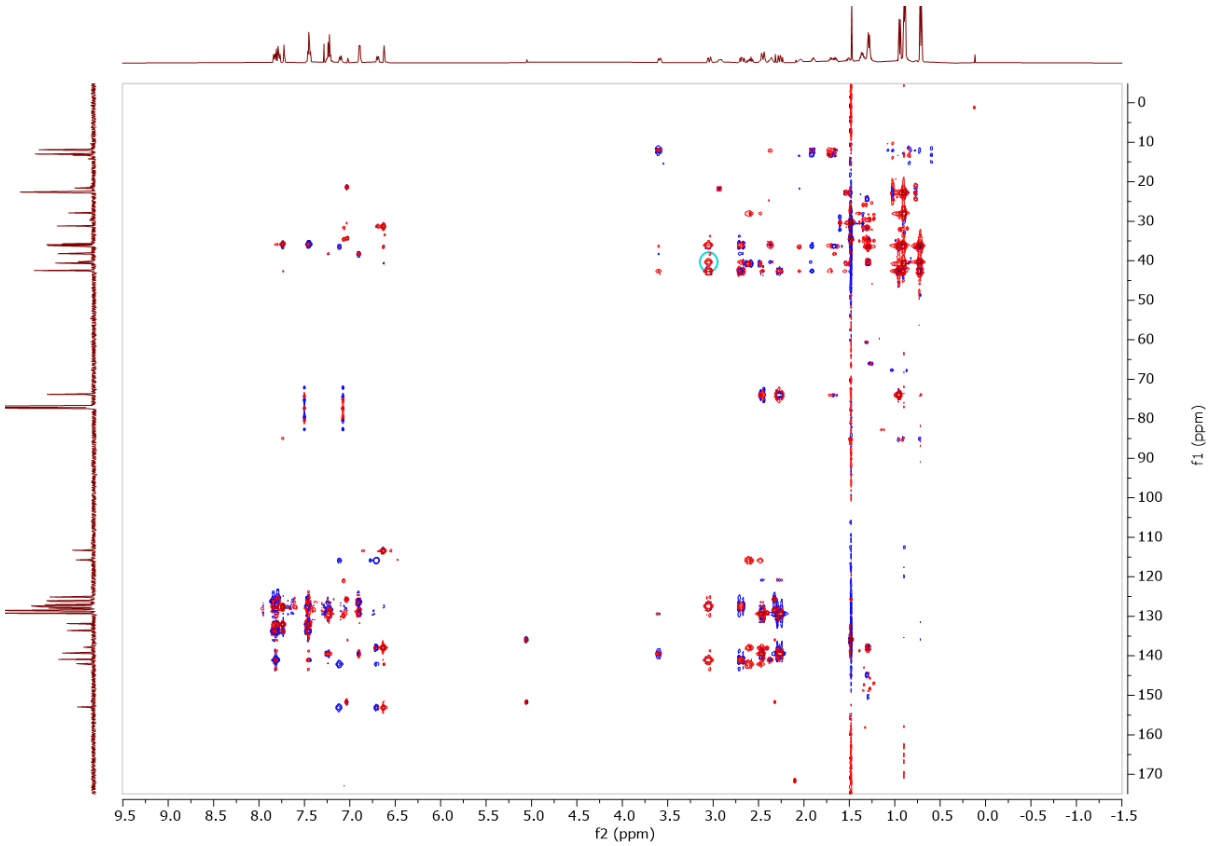


Figure 170. The sum of the IP and AP spectra of the IPAP_HSQMBC of **119.**

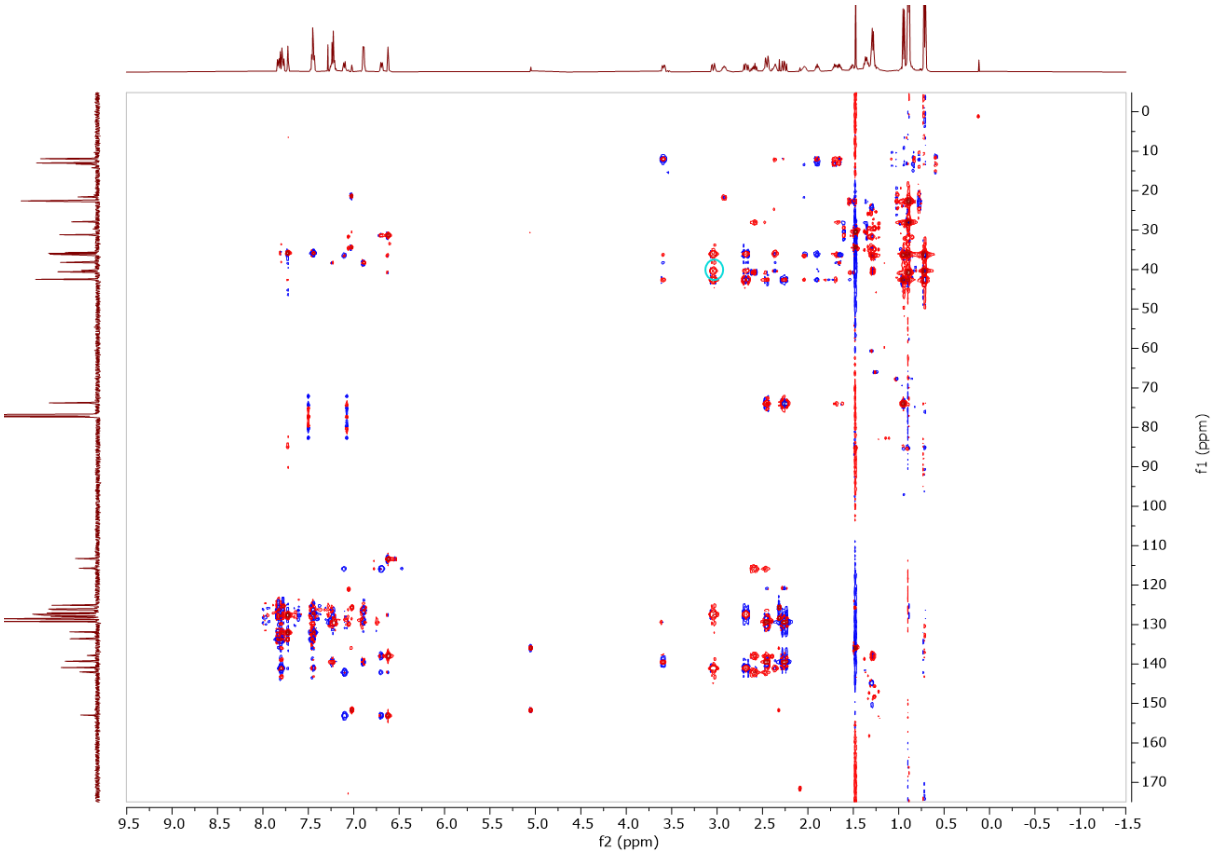


Figure 171. The difference of the IP and AP spectra of the IPAP_HSQMBC of **119.**

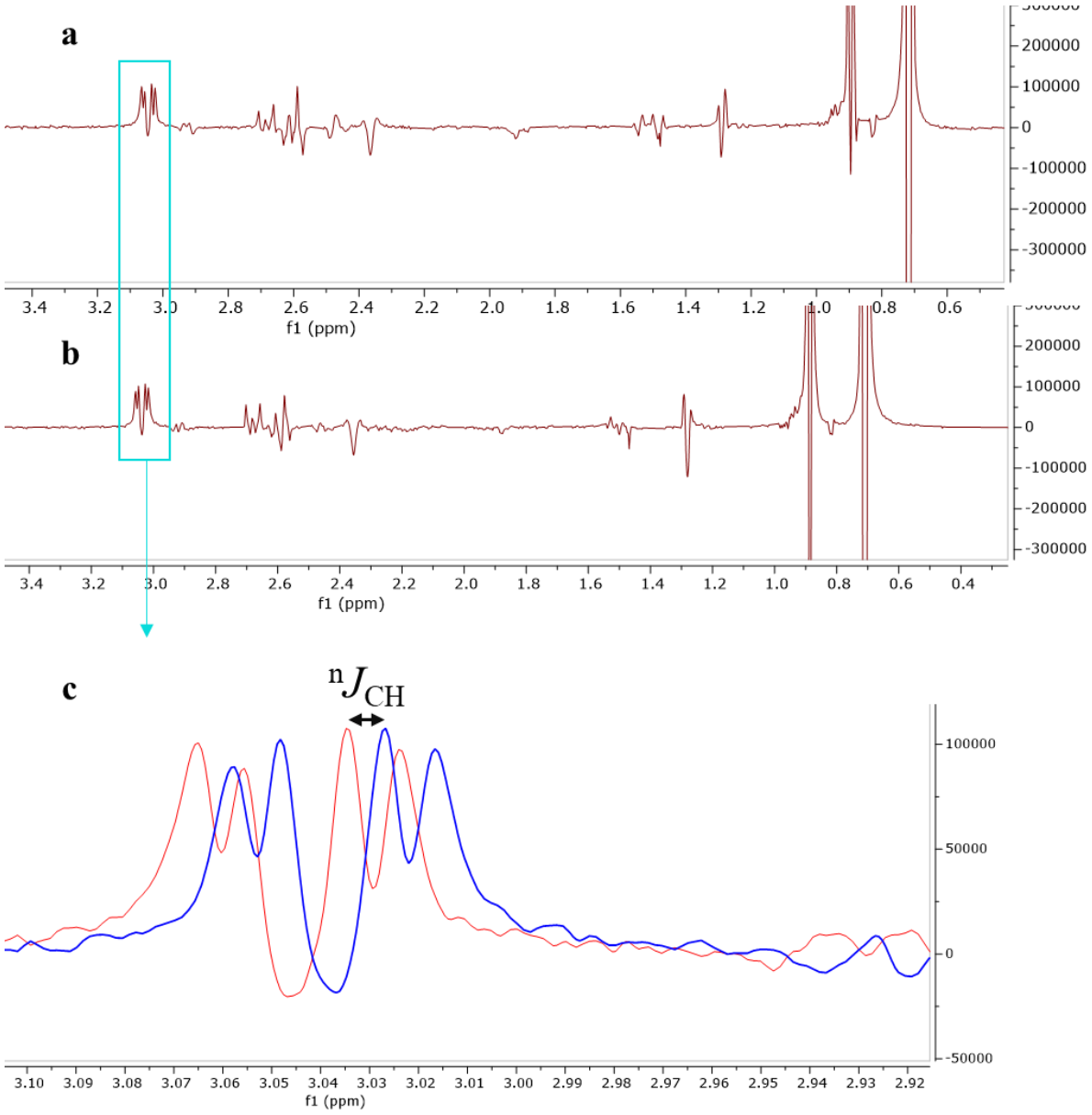


Figure 172. A horizontal 1D trace at 40.3 ppm of both the sum (**a**) and difference (**b**) spectra. **c**) shows the overlay of both the sum and difference 1D-traces for the peak at 3.04 ppm of molecule **119**. The offset between the multiplets is the ^{1}H-^{13}C coupling constant.

9.5.3 The Measurement of Quantitative Interproton Distances

The interproton distances were measured using the CSSF-NOESY experiment (256 scans, 2 s relaxation delay, 500 ms mixing time, 30 ppm (15015 Hz) spectral width, 65536 f1 data points). The data was measured using a Bruker AVANCE III HD 500 MHz NMR spectrometer with a 5 mm DCH ^{13}C-$^{1}H/D$ Cryo Probe.

NOE intensities can be used to derive interproton distances because the distance between two atoms is proportional to the NOE intensity between the same atoms. Thus using a known distance as a reference, typically a methylene pair, it is possible to calculate the interproton distance between two atoms using **Equation 21**.

$$r_{NOE} = r_{ref}\left[\frac{\eta_{NOE}}{\eta_{ref}}\right]^{-1/6}$$

Equation 21. The calculation of interproton distances from NOE intensities.

Where r_{NOE} is the interproton distance between the atoms of interest, r_{ref} is the interproton distance between two reference atoms, η_{NOE} is the NOE intensity between two atoms of interest and η_{ref} is the NOE intensity between two reference atoms. In order to compare all NOE intensities across different 1D-NOESY spectra, differing rates of external relaxation for each proton needs to be corrected for. This can be achieved by applying the PANIC method described by Macura.[106] By setting the irradiated peak to 1000 in all selective 1D-NOSY experiments, differing rates of external relaxation are corrected for.

For **119**, the reference distance was that between $H_{10''}$ and $H_{10'}$ (1.74 Å). The reference distance is incrementally changed to reduce the overall MAD for the dataset. This means that there may be a small %deviation between the experimental reference distance and its calculated value, however the overall MAD will be reduced. Thus, using this reference distance and NOE intensity, all measured NOE intensities can be used to derive an interproton distance (**Figure 173**).

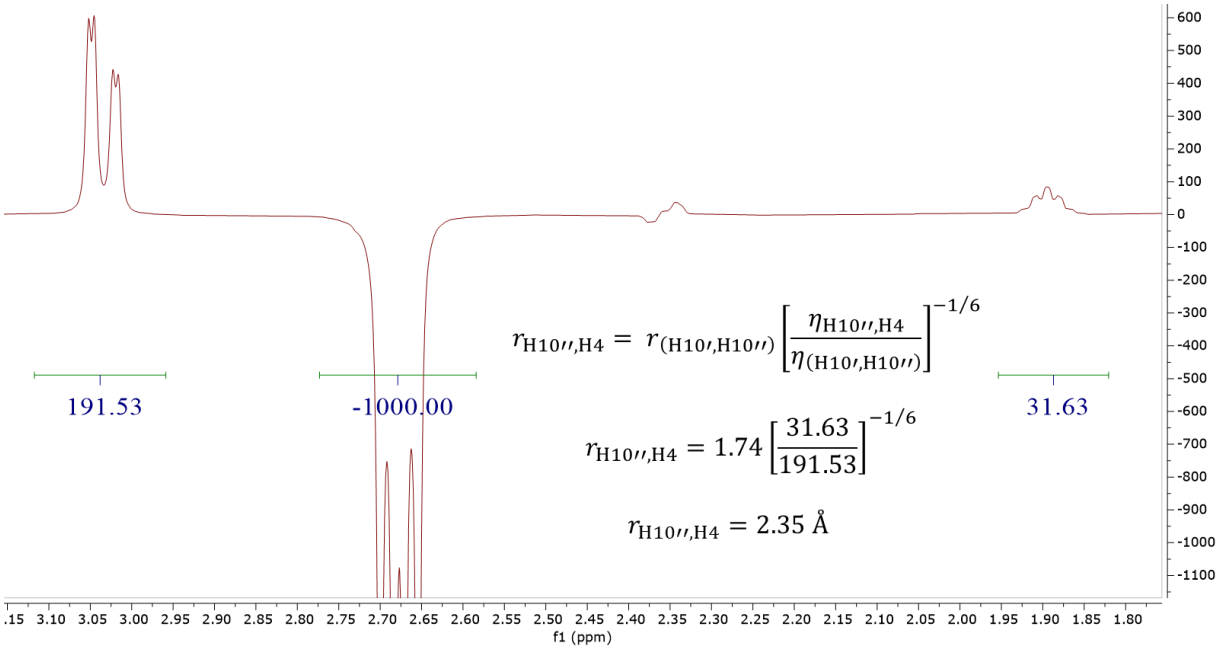


Figure 173. Deriving the $H_{10''}$-H_4 interproton distance.

9.6 The Conformational Analysis of 119 using the hybrid NMR Spectroscopy and Computational Approach

The Boltzmann averaged NMR parameters, calculated as described in the previous section, were then compared to the experimentally measured NMR parameters. The quality of the fit between the two sets of data was assessed using the mean absolute deviation (MAD) and Standard Deviation (SD).

The MAD was calculated using either **Equation 22** when calculating the %deviation between experimental and calculated interproton distances or **Equation 23** when calculating the deviation between experimental and calculated scalar coupling constants and chemical shifts.

$$\%\text{Deviation} = \left(\frac{r_{calc} - r_{exp}}{r_{calc}}\right) x\ 100$$

Equation 22. The calculation of %deviation for interproton distances

$$\text{Deviation} = \left(\delta/J_{calc} - \delta/J_{exp}\right)$$

Equation 23. The calculation of deviation for scalar couplings or chemical shifts.

9.6.1 The Comparison of Experimental and Calculated $^{n}J_{HH}$ Coupling Constants

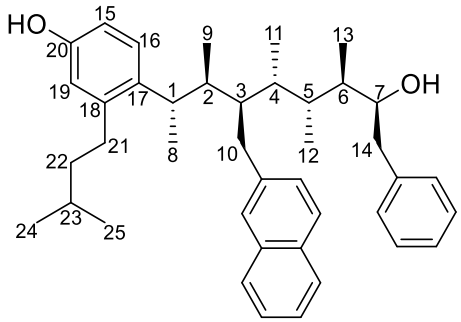


Table 44. A comparison of the experimentally derived $^{n}J_{HH}$ scalar coupling constants to those calculated by DFT calculations for **119**.

Proton	Proton	Experimental Coupling (Hz)	Calculated Coupling (Hz)*	Deviation (Hz)
1	2	10.3	10.7	0.4
1	8	6.7	7.2	0.5
2	3	2.0	2.6	0.6
2	9	6.6	7.4	0.8
3	4	8.2	9.3	1.1
3	10’	3.1	1.7	-1.4
3	10’’	8.4	9.7	1.3
4	5	2.9	3.2	0.3
4	11	6.9	7.4	0.5
5	6	9.2	9.4	0.2
5	12	6.8	7.2	0.4
6	7	4.6	4.9	0.3
6	13	6.8	7.2	0.4
7	14’	9.8	9.1	-0.7
7	14’’	2.2	2.1	-0.1
10’	10’’	15.1	17.6	2.5
14’	14’’	13.4	12.7	-0.7
			Mean Absolute Deviation (Hz)	**0.71**
			Standard Deviation (Hz)	**0.85**

*mPW1PW91/6-311g(d,p)

9.6.2 The Comparison of Experimental and Calculated $^{n}J_{CH}$ Coupling Constants

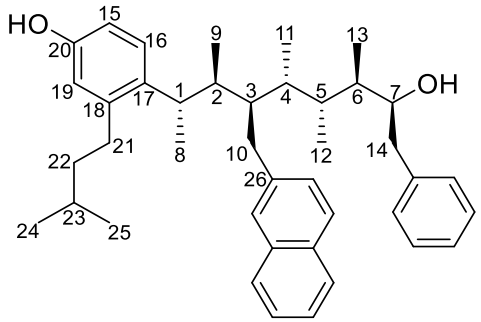


Table 45. A comparison of the experimentally derived $^{n}J_{CH}$ scalar coupling constants to those calculated by DFT calculations for **119**.

Proton	Carbon	Experimental Coupling (Hz)	Calculated Coupling (Hz)*	Deviation (Hz)
1	9	2.5	1.7	-0.7
1	8	4.5	4.3	-0.2
2	8	3.3	3.3	0.0
2	1	5.4	5.3	-0.1
3	11	3.2	2.4	-0.8
3	9	6.1	6.2	0.1
3	4	4.8	5.0	0.2
3	2	4.9	4.7	-0.2
3	26	4.3	4.8	0.4
4	12	5.6	5.5	-0.1
4	2	2.7	2.2	-0.5
4	11	4.7	4.3	-0.4
4	5	5.5	4.7	-0.8
5	7	2.2	2.4	0.2
5	12	4.5	4.3	-0.2
6	12	2.6	2.6	0.1
6	14	5.0	5.2	0.2
6	13	3.7	3.8	0.1
6	5	5.4	4.7	-0.7
7	5	2.0	1.3	-0.7
7	13	4.4	4.5	0.1
7	6	2.5	2.2	-0.3
10’	4	5.7	6.6	0.9
10’	2	3.7	4.6	1.0
10’	3	4.4	4.7	0.3
10’	26	6.7	8.0	1.3
10’’	4	4.5	5.4	0.8
10’’	3	4.6	4.5	-0.1
10’’	26	5.8	5.4	-0.4
13	5	4.1	4.2	0.1
13	6	3.9	3.8	0.0
13	7	4.3	4.3	0.0
14’	6	2.2	2.2	0.1
14’	7	6.8	6.9	0.1

14''	6	1.3	0.8	-0.5
14''	7	1.9	1.8	-0.1
15	19	4.7	4.9	0.2
15	20	2.7	2.6	0.0
16	1	3.5	3.7	0.2
16	20	10.4	10.3	-0.1
19	21	4.7	5.3	0.7
19	1	1.4	0.6	-0.8
19	15	5.0	5.3	0.3
19	20	3.1	3.9	0.8
21'	23	2.6	2.3	-0.3
21'	22	4.9	5.0	0.1
21'	19	5.8	6.1	0.3
			Mean Absolute Deviation (Hz)	**0.35**
			Standard Deviation (Hz)	**0.48**

*mPW1PW91/6-311g(d,p)

9.6.3 The Comparison of Experimental and Calculated Interproton Distances

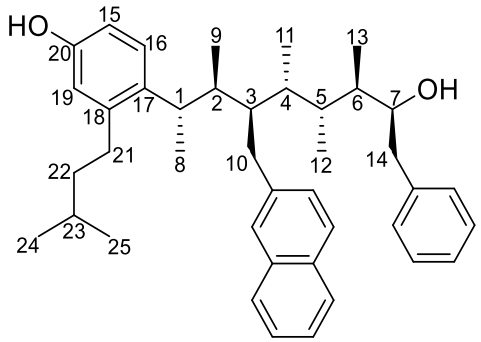


Table 46. A comparison between the experimentally derived interproton distances from 1D-NOE spectroscopy to those calculated by DFT calculations for **119**.

Irradiated ^{1}H (Ha)	Observed ^{1}H (Hb)	Relative η_{AB}	Experimental r_{AB} (Å)	Calculated r_{AB} (Å)*	Deviation (%)
7	14’	17.18	2.60	2.49	-4.42
	14’’	6.77	3.04	3.04	-0.02
	4	14.04	2.69	3.11	15.66
	13	2.37	3.62	3.59	-0.63
	12	12.67	2.74	2.77	1.40
10’	1	60.52	2.11	2.08	-1.48
	10’’	182.70	1.75	1.74	-0.64
	3	19.99	2.54	2.67	5.28
1	10’	34.86	2.31	2.08	-10.13
	20’	40.61	2.25	2.19	-2.84
	20’’	21.76	2.50	2.45	-2.20
	3	11.73	2.77	2.91	5.03
	2	5.66	3.13	3.02	-3.56
10’’	**10’ (r_{ref})**	**191.53**	**1.74**	**1.74**	0.15
	4	31.63	2.35	2.41	2.53
3	10’	15.69	2.64	2.67	1.11
	1	13.11	2.72	2.91	7.00
	2	25.06	2.44	2.48	1.54
	4	13.14	2.72	2.83	4.08
	8	18.88	2.56	2.51	-1.97
	11	4.27	3.28	3.27	-0.19
2	16	48.81	2.19	2.27	4.10
	3	22.77	2.48	2.48	-0.07
	8	6.96	3.02	2.93	-3.18
	11	15.60	2.64	2.68	1.46
4	7	14.60	2.67	3.11	16.42
	10’’	20.57	2.52	2.41	-4.56
	3	11.13	2.80	2.83	1.24
	13	15.17	2.66	2.75	3.49
				Mean Absolute Deviation (%)	**3.67**
				Standard Deviation (%)	**5.42**

*mPW1PW91/6-311g(d,p)

9.6.4 The Comparison of Experimental and Calculated Chemical Shifts

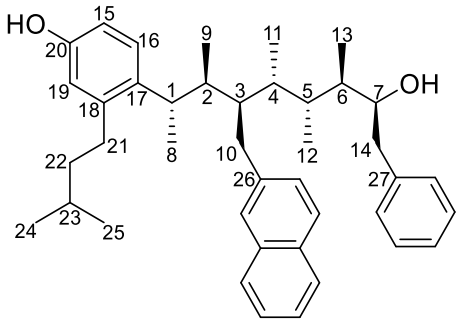


Table 47. A comparison of the experimental δ_H to those calculated by DFT calculations for **119**.

Proton	Experimental δ_H (ppm)	Magnetic Shielding Tensor*	Calculated δ_H (ppm)	Scaled δ_H (ppm)	Deviation (ppm)
1	2.91	28.80	3.13	3.00	0.09
2	2.03	29.79	2.14	2.05	0.02
3	2.35	29.34	2.59	2.48	0.13
4	1.88	29.84	2.08	2.00	0.12
5	1.70	30.23	1.69	1.62	-0.08
6	1.64	30.21	1.71	1.64	0.00
7	3.57	28.67	3.25	3.12	-0.45
8	1.27	30.39	1.54	1.47	0.20
9	0.70	31.23	0.69	0.66	-0.04
10’	3.02	28.54	3.38	3.24	0.22
10’’	2.67	29.28	2.65	2.54	-0.13
11	0.88	31.03	0.90	0.86	-0.02
12	0.70	31.44	0.49	0.47	-0.23
13	0.93	30.89	1.04	1.00	0.07
14’	2.24	29.66	2.26	2.17	-0.07
14’’	2.45	29.70	2.23	2.14	-0.31
15	6.68	25.08	6.85	6.57	-0.11
16	7.09	24.43	7.50	7.19	0.10
19	6.60	24.95	6.98	6.70	-0.10
21’	2.59	29.03	2.90	2.78	0.19
21’’	2.45	29.36	2.56	2.46	0.01
22’	1.34	30.43	1.49	1.43	0.09
22’’	1.34	30.50	1.43	1.37	0.03
23	1.50	30.30	1.63	1.56	0.06
24/25	0.87	31.00	0.93	0.89	0.02
			Mean Absolute Deviation (ppm)		**0.11**
			Standard Deviation (ppm)		**0.16**

*mPW1PW91/6-311g(d,p)

Table 48. A comparison of the experimental δ_c to those calculated by DFT calculations for **119**.

Carbon	Experimental δ_C (ppm)	Magnetic Shielding Tensor*	Calculated δ_C (ppm)	Scaled δ_C (ppm)	Deviation (ppm)
1	36.5	149.09	40.3	37.0	0.5
2	40.3	144.69	44.7	41.3	1.0
3	42.6	143.36	46.0	42.5	-0.1
4	36.0	148.10	41.3	38.0	2.0
5	36.3	149.64	39.7	36.5	0.2
6	42.6	144.86	44.5	41.1	-1.5
7	74.0	113.05	76.3	71.7	-2.3
8	21.7	164.82	24.5	21.9	0.2
9	13.4	175.48	13.9	11.6	-1.8
10	35.8	148.60	40.8	37.5	1.7
11	12.1	175.66	13.7	11.4	-0.7
12	13.1	175.38	14.0	11.7	-1.4
13	12.0	175.53	13.8	11.6	-0.4
14	38.3	147.95	41.4	38.1	-0.2
15	113.4	73.01	116.4	110.2	-3.2
16	127.5	55.28	134.1	127.3	-0.2
17	138.0	45.53	143.9	136.7	-1.3
18	142.1	39.14	150.2	142.8	0.7
19	115.8	69.17	120.2	113.9	-1.9
20	153.1	28.42	161.0	153.2	0.1
21	31.3	152.29	37.1	33.9	2.6
22	40.7	144.64	44.7	41.3	0.6
23	28.0	155.76	33.6	30.6	2.6
24/25	22.7/22.7	164.47/165.67	24.9/23.7	22.2/21.1	-0.5/-1.6
26	141.0	39.55	149.8	142.5	1.5
27	139.4	42.42	147.0	139.7	0.3
				Mean Absolute Deviation (ppm)	**1.14**
				Standard Deviation (ppm)	**1.46**

*mPW1PW91/6-311g(d,p)

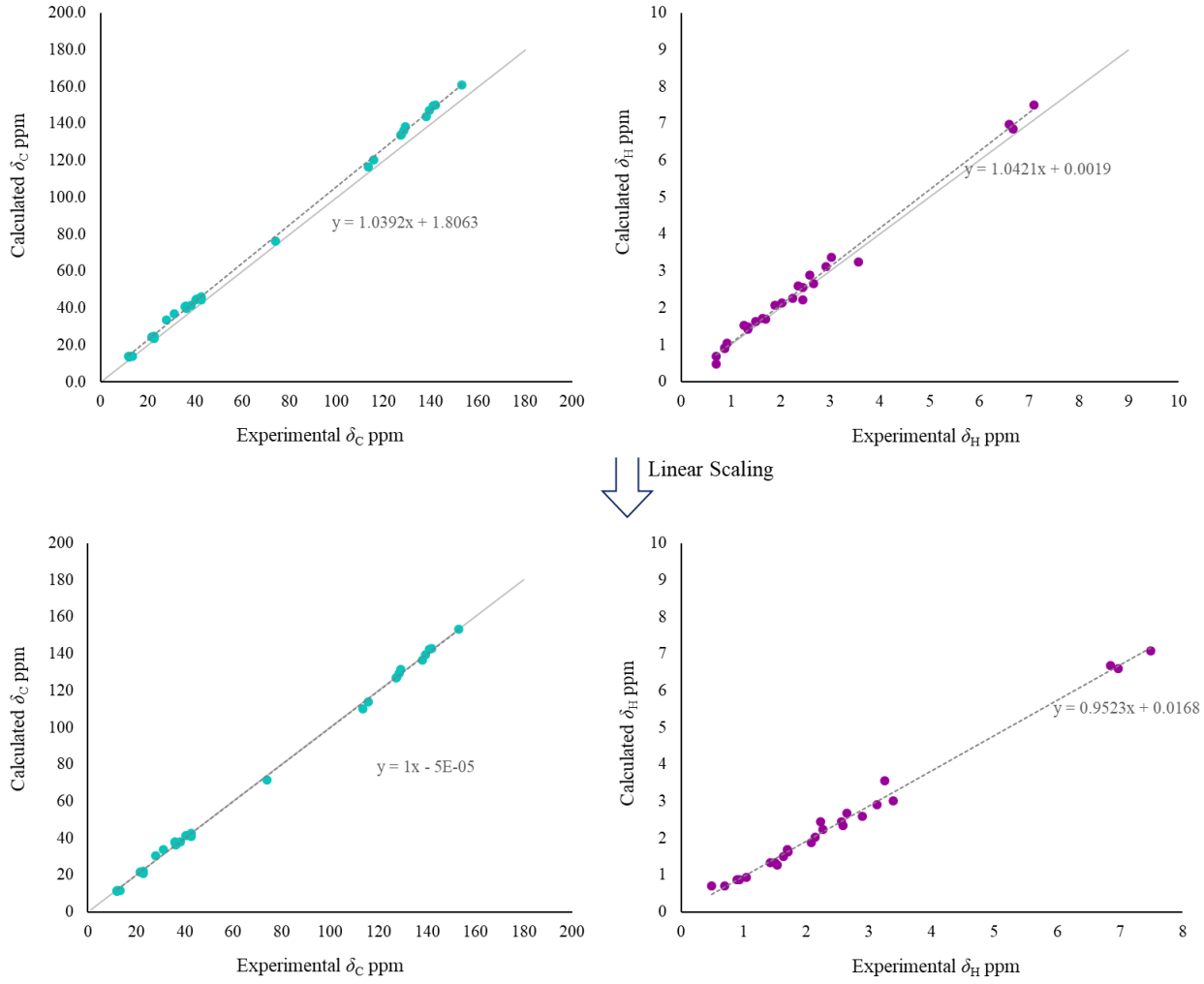


Figure 174. The linear scaling applied to the calculated chemical shifts for **119**.

9.7 The Binding of the Designed Inhibitors to Mdm2 using ^{1}H-^{15}N HSQC Spectroscopy

9.7.1 The Expression and Purification of ^{15}N-Mdm2

Sequence Used:

MHHHHHHGKPIPNPLLGLDSTENLYFQGIDPFTLVRPKPLLLKLLKSVGAQKDTYTMKEVLF
YLGQYIMTKRLYDEKQQHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIYRNLVVVNQQESSD
SGTSVSEN

A fragment of the N-terminal domain of human Mdm2 (residues 27-127) was cloned into the pET-20 (Novagen) vector and expressed in the *E.Coli* BL21(DE3) strain. 200 µL of sterile media was added to the BL21(DE3) cells and incubated for one hour at 37 °C. *E.Coli* BL21(DE3) cells were grown on an LB agar plate overnight at 37 °C. Three cell colonies were harvested and added to 50 mL of LB media. 50 µL of a 50 mg/mL solution of ampicillin was added for plasma selection. Cell colonies were incubated at 37 °C overnight. 2 mL of the seed broth was used to inoculate flasks containing 200 mL of LB media and 400 µL of ampicillin. Flasks were incubated at 37 °C until an OD_{600} of 1.0-1.4 was reached. Cell cultures were cooled to 16 °C and induced with 200 µM of IPTG. Protein expression was maintained overnight at 16 °C. Cells were harvested by centrifugation at 6000 rpm at 4 °C. The cell pellet was resuspended in 20 mL of column buffer (50 mM Tris base, 0.5 M NaCl, pH 8.0) followed by flash freezing in liquid nitrogen. The cell suspension was stored at -4 °C until ready for purification. Prior to purification, the cells were lysed at -10 °C by sonication for a total time of 10 minutes consisting of cycles of 5 seconds sonication followed by 20 seconds rest. Cells were clarified by centrifuged at 15000 rpm and the cell lysate was purified using a HisTrap HP 5 mL Ni column. Protein was eluted via a linear gradient from 5 to 100% elution buffer (50 mM Tris buffer at pH 8.0 with 0.3 M NaCl and 0.8 M imidazole) and collected in 1 mL fractions. The pure fractions were analysed by SDS PAGE (**Figure 175**) and those containing the protein of interest were combined.

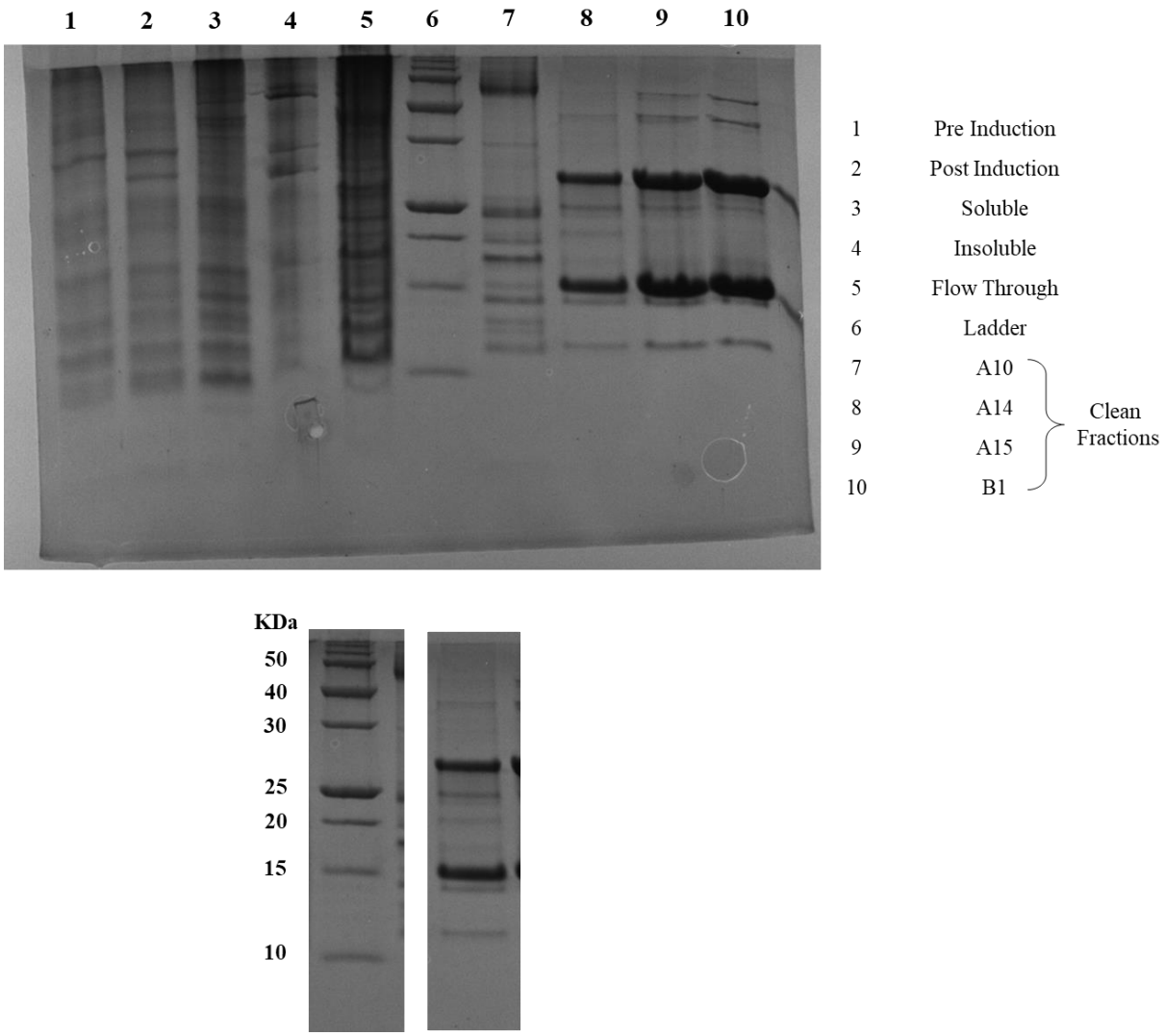


Figure 175. The SDS-PAGE results following the first purification using a HisTrap HP 5 mL Ni column.

TEV protease was used to attempt to cleave the His_6-tag however, all conditions used were unsuccessful. The combined fractions after the first HisTrap HP 5 mL Ni column were subjected to a second purification using a HiLoad 26/60 Superdex 75 prep grade column, equilibrated with 20 mM Tris, 250 mM NaCl, 1 mM TCEP at pH 7.5. SDS-PAGE was used to confirm the fractions containing pure Mdm2 (**Figure 176**).

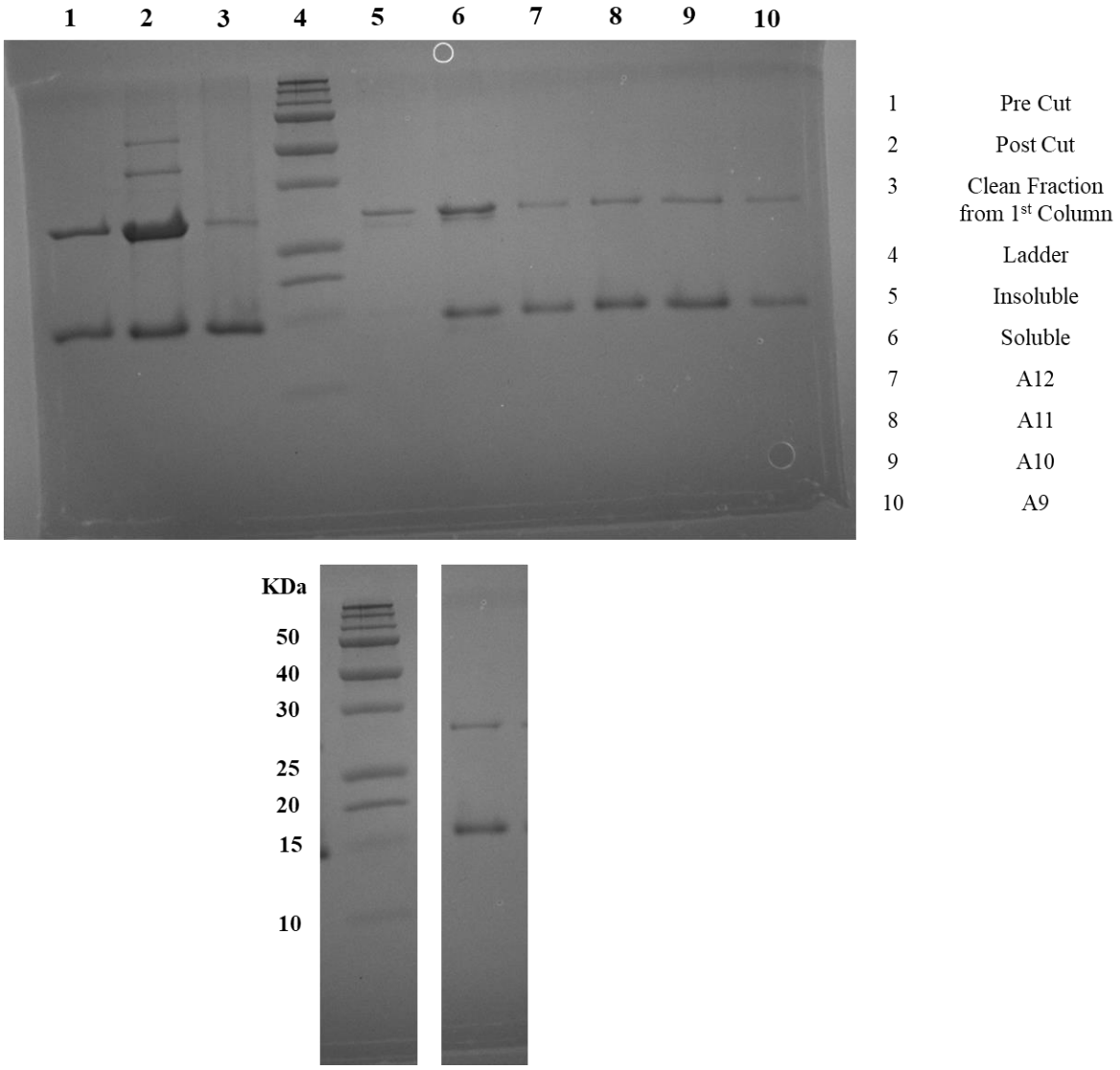


Figure 176. The SDS-PAGE results following the second purification using a HiLoad 26/60 Superdex 75 prep grade column.

Samples were desalted for MALDI mass spectrometry using a C_4 ZipTip^{TM} (Millipore). The ZipTip was activated with 3 10 µL washes of 50% acetonitrile in water and equilibrated with five 10 µL washes of 0.1% TFA in water. A 10 µL sample was then loaded onto the ZipTip with 15 washes. The protein was desalted with ten 10 µL washes with 5% MeOH, 0.1% TFA in water and discarded to waste. Finally, the sample was eluted with 10 µL of 70% acetonitrile, 0.1% TFA in water. MALDI mass spectrometry confirm that Mdm2 had been successfully expressed, and a M_w of 15233.0 Da was observed, corresponding to the expected average M_w of the His_6-tagged Mdm2 (sequence shown below). Obs: 15233.00 Da Expt: 15232.51 Da (**Figure 177**).

MHHHHHHGKPIPNPLLGLDSTENLYFQGIDPFTLVRPKPLLLKLLKSVGAQKDTYTMKEVLF
YLGQYIMTKRLYDEKQQHIVYCSNDLLGDLFGVPSFSVKEHRKIYTMIYRNLVVVNQQESSD
SGTSVSEN

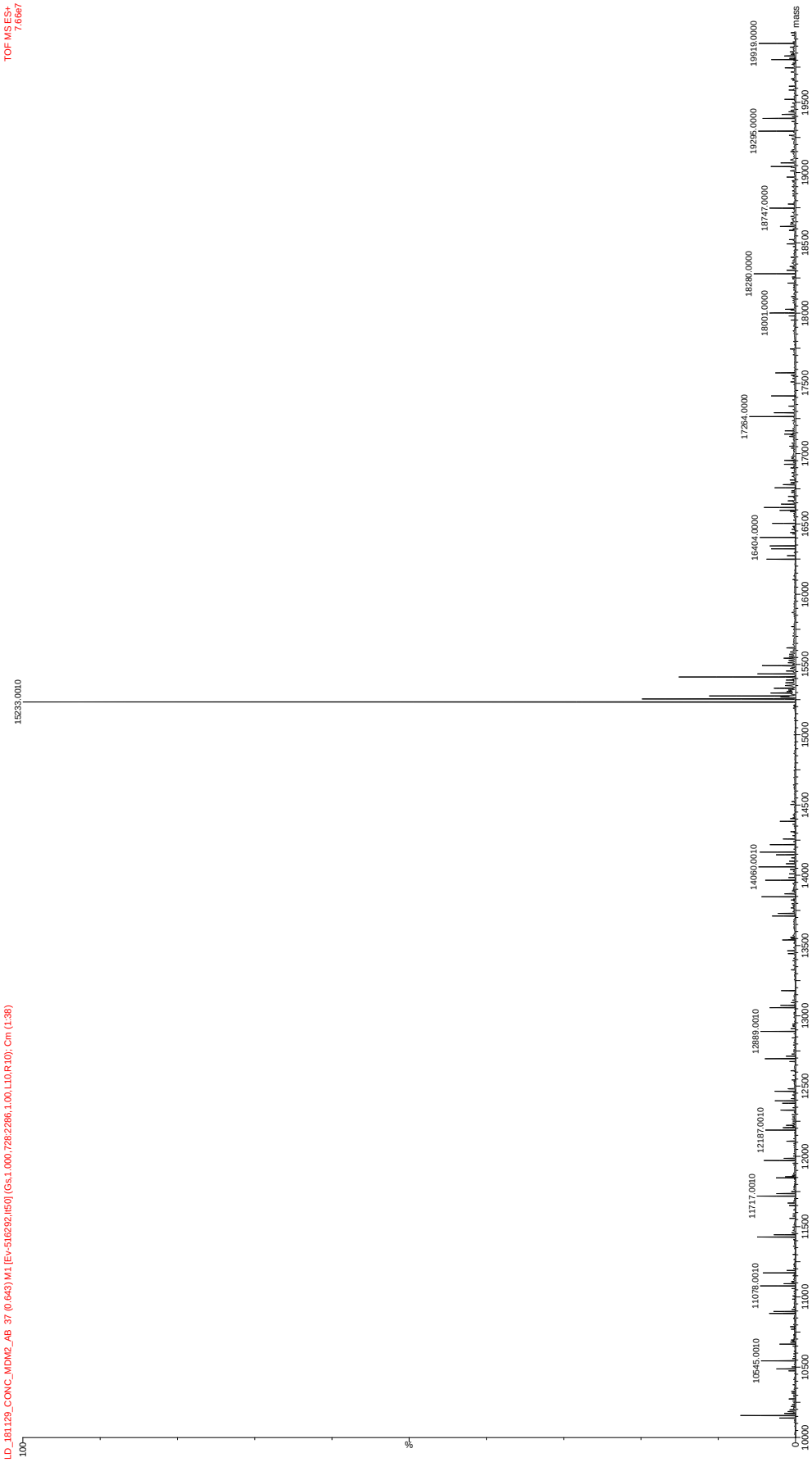


Figure 177. The results from the MALDI mass spectroscopy of Mdm2, showing the Mw found.

Following the successful expression of unlabelled Mdm2, ^{5}N-labelled Mdm2 was expressed. Three colonies of BL21-(DE3) cells were harvested and added to 50 mL of LB media. 50 µL of a 50 mg/mL solution of ampicillin was added for plasma selection. Cell colonies were incubated at 37 °C overnight. 2 mL of the seed broth was used to inoculate flasks containing 200 mL of LB media and 400 µL of ampicillin. Flasks were incubated at 37 °C until an OD_{600} of 1.0-1.4 was reached. Cells were harvested by centrifugation at 6000 rpm and washed twice with sterile M9 media. The cells were resuspended in M9 and inoculated with 450 µL of ampicillin. $^{15}NH_4Cl$ (1 g/L), glucose (3 g/L), $MgSO_4$ (200 µg/mL) and a trace metal mix (150 µL) were added followed by incubation at 37 °C for 1 hour. The cultures were induced with IPTG (200 µM) at 16 °C overnight. Cells were harvested by centrifugation at 6000 rpm and purified as described for the unlabelled Mdm2. MALDI mass spectroscopy confirmed that full ^{15}N labelling had occurred. Obs: 15411.0 Expt: 15414.1 (**Figure 178**)

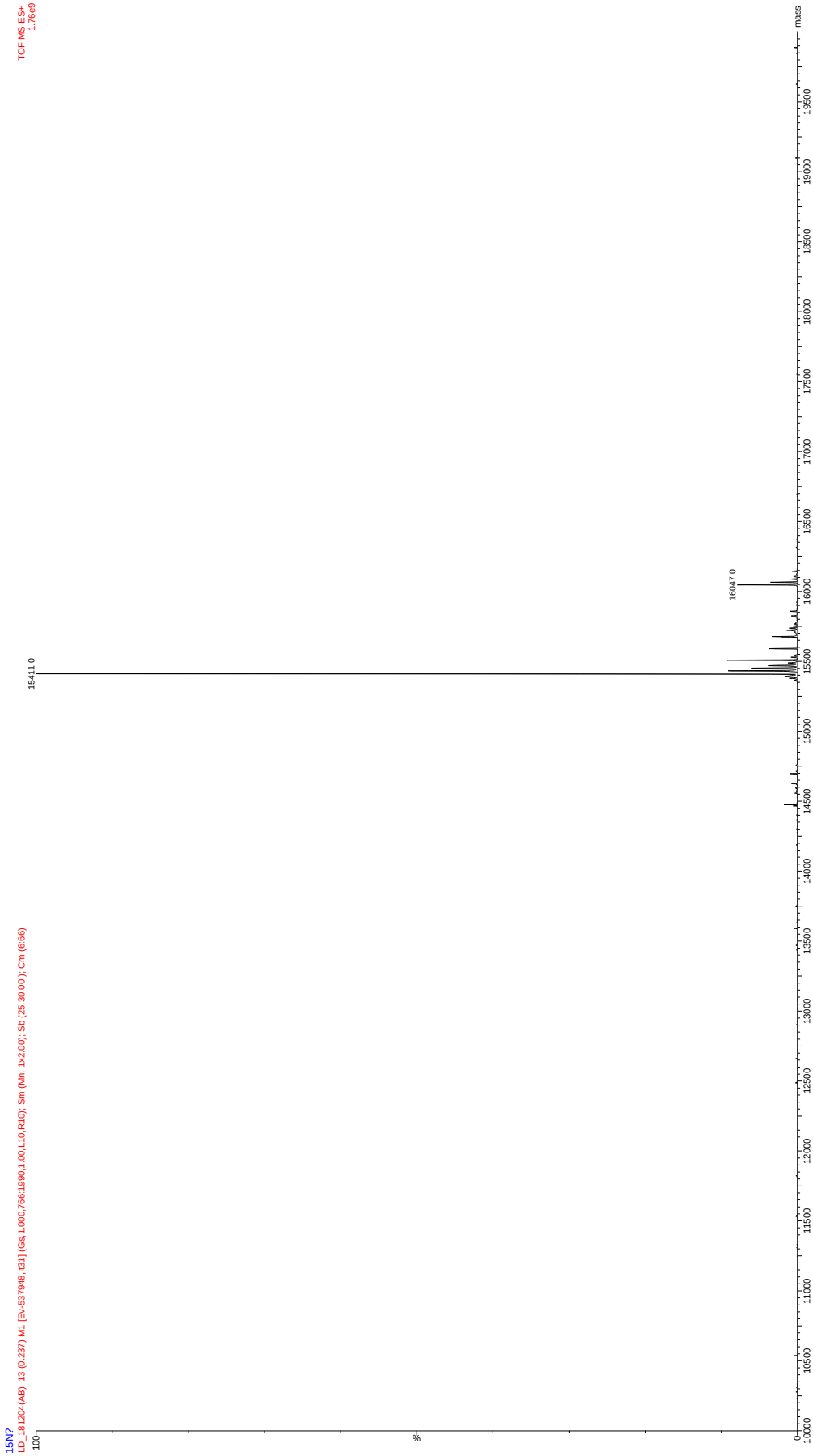


Figure 178. The results from the MALDI mass spectroscopy of ^{15}N-Mdm2, showing the M_w found.

NMR spectroscopy was performed to confirm the purity of Mdm2 and that it had folded well. A 2D-^{1}H-^{15}N TROSY of Mdm2 was acquired at a concentration of 130 μM on a 700 MHz spectrometer fitted with a 1.7 mm inverse triple resonance micro-Cryo Probe at 20 °C. The DPFGSE sequence was used for water suppression. The protein was solubilised in a 10 mM $NaHPO_4$ solution with 150 mM NaCl at pH 7.4. 10% D_2O was added to the sample to provide a lock signal. The total sample volume was 50 μl using 1.7 mm NMR tubes. The acquisition parameters include: 512 scans, 256 f1 increments and 512 f2 data points, f1 spectral width of 25 ppm (1774 Hz), f2 spectral width of 12 ppm (8403 Hz). Spectra were processed and analysed in MestreNova v14.0. Spectra were zero filled to 4096 data points and manually phase corrected using the largest peak as the pivot point. The acquired TROSY was partially assigned using previously reported assignments of Mdm2 in the literature (**Figure 92**, **Table 49**).[300,306]

Table 49. A comprehensive list of the resolved peaks with their assignments

F1 (ppm)	F2 (ppm)	Assignment
8.20	118.35	14Thr
7.51	116.50	26Thr
8.45	128.15	27Leu
9.66	116.10	28Val
9.25	121.99	29Arg
8.71	118.64	33Leu
7.30	118.35	34Leu
8.56	120.22	(35Leu/56Tyr)
7.35	121.33	37Leu
8.18	120.52	38Leu
8.66	118.33	39Lys
7.83	116.57	40Ser
7.23	112.63	41Val
7.46	106.35	42Gly
7.31	124.78	43Ala
8.70	118.06	44Gln
7.67	116.26	45Lys
6.97	111.17	47Thr
8.59	120.72	48Tyr
8.89	110.11	49Thr
8.25	121.14	50Met
7.82	119.93	52Glu
8.30	118.93	54Leu
8.17	122.73	55Phe
8.63	121.30	57Leu
8.26	106.37	58Gly
7.82	121.83	59Gln
8.76	124.15	60Tyr
8.58	119.18	61Ile
7.92	116.17	62Met
8.63	116.04	63Thr
7.56	116.07	65Arg

7.61	115.17	66Leu
6.75	113.96	67Tyr
8.49	123.74	68Asp
8.54	124.00	69Glu
8.21	116.67	70Lys
8.29	120.92	71Gln
8.29	119.39	72Gln
7.96	113.50	73His
7.58	122.30	74Ile
8.27	125.89	75Val
8.24	125.73	76Tyr
8.38	117.27	77Cys
7.62	118.11	80Asp
9.81	118.83	82Leu
7.77	106.86	83Gly
7.49	123.04	84Asp
8.00	119.38	85Leu
8.36	116.99	86Phe
8.24	108.39	87Gly
7.05	108.50	88Val
7.21	111.25	90Ser
8.27	114.43	91Phe
8.71	112.86	92Ser
9.37	122.42	93Val
7.97	117.28	94Lys
7.34	120.37	95Glu
7.69	120.48	96His
7.73	121.32	98Lys
7.74	119.05	99Ile
8.39	121.11	100Tyr
8.12	112.73	101Thr
7.59	120.05	102Met
8.73	120.22	104Tyr
8.67	127.85	(108Val/109Val)
8.29	127.29	110Val
8.25	123.42	111Asn
8.37	122.07	117Asp
8.31	115.93	118Ser
7.45	119.18	None
7.49	118.78	None
7.71	118.35	None
7.77	118.46	None
7.80	118.18	None
7.88	117.90	None
7.97	124.49	None
7.99	114.62	None
8.03	115.45	None
8.04	118.96	None
8.05	122.01	None

8.09	118.63	None
8.12	118.45	None
8.14	121.59	None
8.32	116.80	None
8.39	121.42	None
8.40	123.48	None

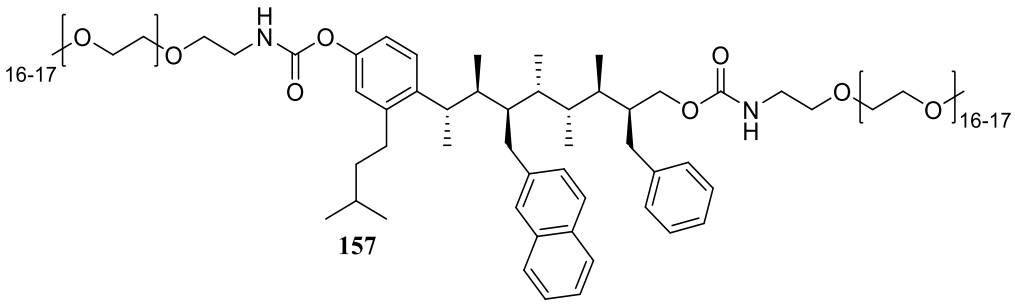
9.7.2 The binding of 157 by TROSY spectroscopy

A 2D-^{1}H-^{15}N TROSY of Mdm2 was acquired at a concentration of 50 μM on a 700 MHz spectrometer fitted with a 1.7 mm inverse triple resonance micro-Cryo Probe at 20 °C. The DPFGSE sequence was used for water suppression. The protein was solubilised in a 20 mM Tris buffer with 250 mM NaCl at pH 7.5. 10% D_2O was added to the sample to provide a lock signal. 10% Acetone-d_6 was added for ligand solubility. The total concentration of ligand **157** was 70 μM. The total sample volume was 50 μl using 1.7 mm NMR tubes. The acquisition parameters include: 512 scans, 256 f1 increments and 512 f2 data points, f1 spectral width of 25 ppm (1774 Hz), f2 spectral width of 12 ppm (8403 Hz). A 2D-^{1}H-^{15}N TROSY of Mdm2 (50 μM) with ligand **157** (70 μM) in the same buffer conditions was acquired with the same acquisition parameters. Spectra were processed and analysed in MestreNova v14.0. Spectra were zero filled to 4096 data points and manually phase corrected using the largest peak as the pivot point. The overlaid spectra of *apo*-Mdm2 and Mdm2+ligand **157** can be seen in **Figure 179**

Chemical shift perturbations (CSP's) were calculated as a Euclidean distance using **Equation 24** with a value of α=0.14. The value of α was calculated from the ratio of the chemical shift range (in ppm) of ^{1}H to that of ^{15}N (3.08/22.1 = 0.139). The calculated CSPs, or Euclidean Distances, are shown in **Table 50**.

$$\delta_{obs} = \sqrt{\frac{1}{2}[\delta_H^2 + (\alpha\delta_N^2)]}$$

Equation 24. The calculation of Euclidean distance



157

Scheme 31. The structure of Ligand **157**

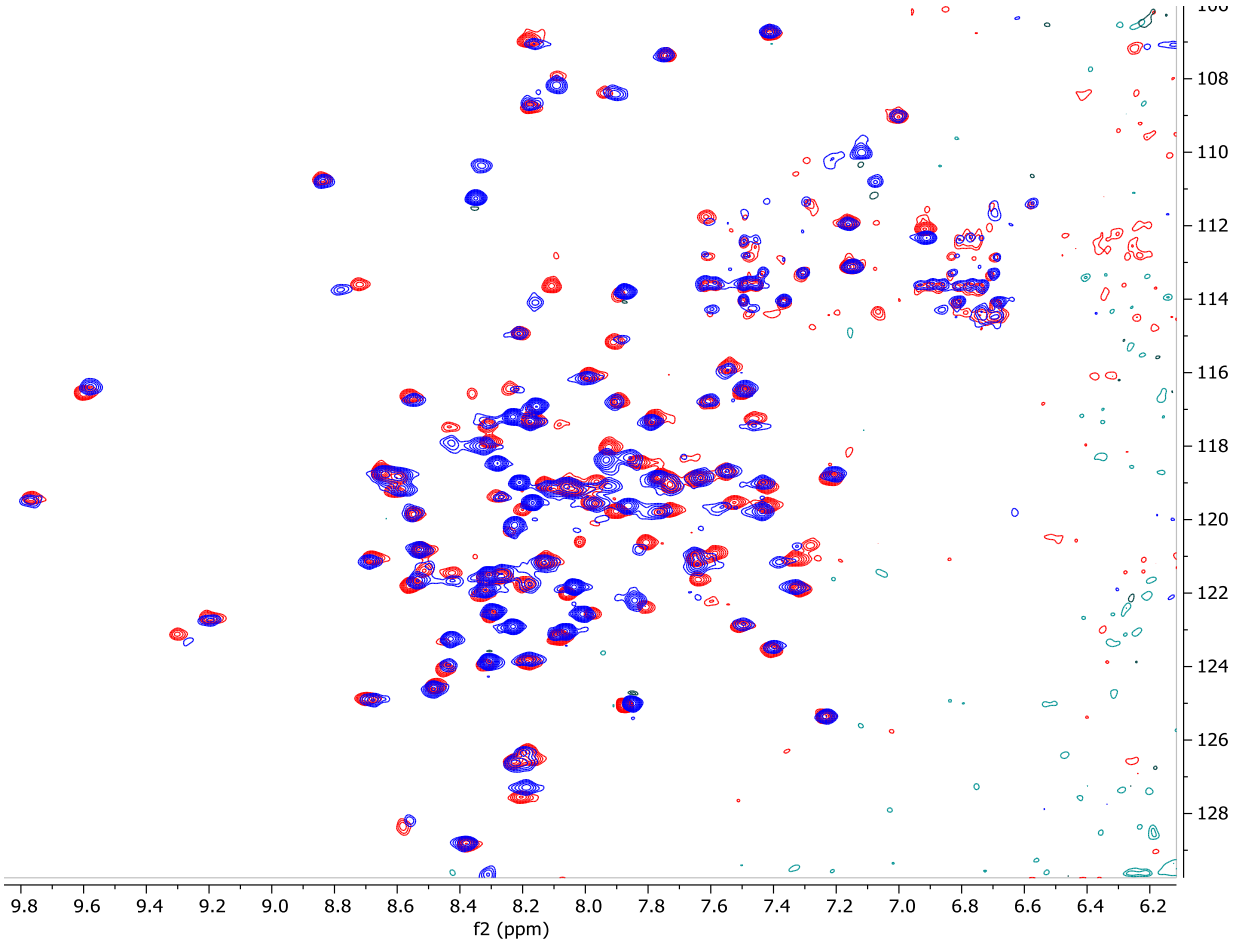


Figure 179. The ^{1}H-^{15}N TROSY of Mdm2 (red) and Mdm2+ligand **157** (blue)

Table 50. The Euclidean Distances for the CSP's observed upon binding of ligand **157** to Mdm2. The first section corresponds to CSPs >2σ, the second section corresponds to CSPs >σ while the last section corresponds to CSPs <σ.

Residue	$\Delta\delta_H$ (ppm)	$\Delta\delta_N$ (ppm)	Euclidean Distance (α = 0.14) (ppm)
T101	0.052	0.450	0.176
F86	0.007	0.470	0.176
K98	0.002	0.420	0.157
K94	0.009	0.350	0.131
V110	0.018	0.270	0.103
T26	0.002	0.250	0.094
T47	0.007	0.240	0.090
V93	0.037	0.210	0.087
S92	0.062	0.160	0.086
Y100	0.003	0.220	0.082
E52	0.023	0.210	0.082
V53	0.033	0.200	0.082
F55	0.025	0.180	0.072
V28	0.021	0.180	0.071
M102	0.042	0.140	0.067
L85	0.038	0.140	0.065
V108	0.023	0.160	0.064
H73	0.011	0.160	0.061
E95	0.052	0.080	0.060
L57	0.024	0.140	0.058
D68	0.011	0.140	0.054
C77	0.010	0.140	0.053
L54	0.022	0.120	0.050
L34	0.015	0.120	0.047
T63	0.015	0.120	0.047
K51	0.022	0.111	0.047
Y104	0.012	0.120	0.046
S40	0.007	0.120	0.045
L66	0.004	0.120	0.045
Y56	0.017	0.110	0.045
G58	0.021	0.100	0.043
I99	0.037	0.051	0.042
G87	0.002	0.110	0.041
F91	0.029	0.078	0.041
R65	0.008	0.100	0.038
E69	0.004	0.090	0.034
D111	0.015	0.080	0.033
Q44	0.012	0.080	0.032
L33	0.023	0.060	0.032
D84	0.011	0.080	0.032
I74	0.014	0.070	0.030
Y60	0.025	0.040	0.029
S118	0.023	0.040	0.027
L37	0.015	0.050	0.024
M62	0.018	0.040	0.023
M50	0.021	0.010	0.021
I61	0.017	0.020	0.019
	Standard Deviation (ppm)		**0.037**

9.7.3 The Estimation of K_d using ^{1}H-^{15}N SOFAST HMQC for 157

A 2D ^{1}H-^{15}N SOFAST HMQC of Mdm2 at a concentration of 70 μM with five concentrations of ligand **157** ranging from 17.5 to 280 μM was acquired on a 700 MHz spectrometer fitted with a 1.7 mm inverse triple resonance micro-Cryo Probe 20 °C using the sofasthmqc sequence. The DPFGSE sequence was used for water suppression. The protein was solubilised in a 20 mM Tris buffer with 250 mM NaCl at pH 7.5. 10% D_2O was added to the sample to provide a lock signal. 10% Acetone-d_6 was added for ligand solubility. A 10 mM stock solution of ligand **157** in acetone-d_6 was prepared and the required volume was added to the protein sample to give final ligand concentrations of 17.5 μM, 35 μM, 70 μM, 140 μM and 280 μM. The total sample volume was 50 μL using 1.7 mm NMR tubes. The acquisition parameters include: 64 scans, 128 f1 increments and 757 f2 data points, f1 spectral width of 25 ppm (1774 Hz), f2 spectral width of 12 ppm (8418 Hz). Spectra were processed and analysed in MestreNova v14.0. Spectra were zero filled to 4096 data points and manually phase corrected using the largest peak as the pivot point. **Figure 180** shows an overlay of the spectrum obtained from all five titration points and *apo*-Mdm2.

The ^{1}H-^{15}N SOFAST HMQC spectra were analysed using Mbinding from MestreNova. The Euclidean distances are calculated using **Equation 24** and the value of α used was 0.14. The experimental CSP's were used to calculate binding curves. The CSP's measured at different protein concentrations can be used in a nonlinear least-square fitting to calculate K_d using **Equation 25**. The final K_d was calculated as the average of the individual values of K_d obtained for 20 different peaks. A summary of the K_d obtained for each peak is provided in **Table 51**. Example binding data obtained using Mbinding for three different peaks is also shown below.

$$\Delta\delta_{obs} = \Delta\delta_{max}\frac{\left\{([P]_t + [L]_t + K_d) - [([P]_t + [L]_t + K_d)^2 - 4[P]_t[L]_t]^{\frac{1}{2}}\right\}}{2[P]_t}$$

Equation 25. Using CSP's to estimate K_d.

Where $\Delta\delta_{obs}$ is the change in the observed shift from the shift of the native protein, $\Delta\delta_{max}$ is the maximum chemical shift change on saturation of the protein with ligand, $[P]_t$ is the total protein concentration and $[L]_t$ is the total ligand concentration.

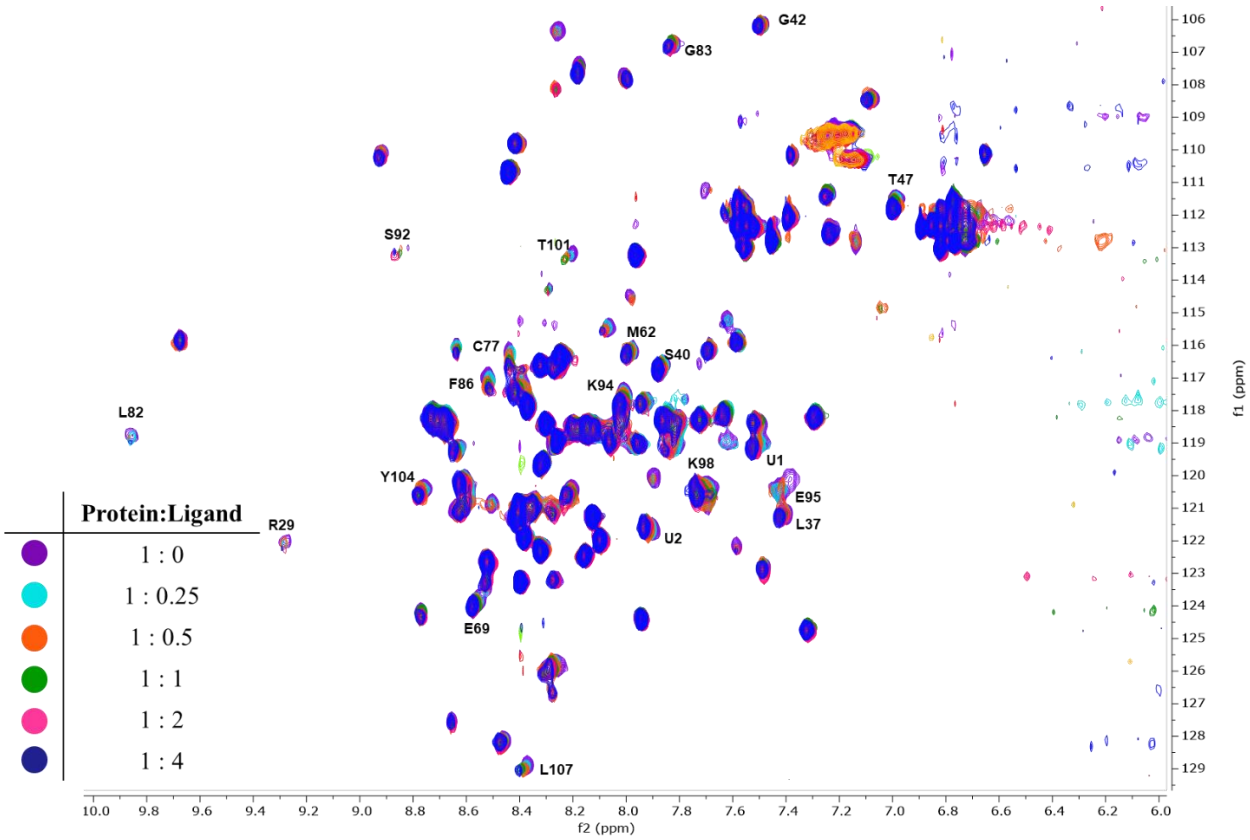


Figure 180. The titration of ligand **157** against Mdm2. Purple (70 μM Mdm2), light blue (70 μM Mdm2 + 17.5 μM ligand), orange (70 μM Mdm2 + 35 μM ligand), green (70 μM Mdm2 + 70 μM ligand), pink (70 μM Mdm2 + 140 μM ligand), dark blue (70 μM Mdm2 + 280 μM ligand). The 20 peaks that were considered in the calculation of K_d are highlighted.

Table 51. The calculated value of K_d and maximum CSP for 20 different residues interacting with **157**. The average K_d and σ K_d are also shown.

Residue	K_d (μM)	CSP Max (ppm)
C77	16.91	0.09
E95	15.98	0.03
F86	13.21	0.06
G42	20.28	0.02
G83	24.96	0.03
K94	12.27	0.06
K98	12.90	0.05
L107	15.17	0.04
L37	23.39	0.07
L82	21.47	0.04
M62	20.88	0.03
R29	17.07	0.05
S40	18.17	0.03
S92	11.62	0.06
T101	10.49	0.07
T47	12.08	0.05
Y104	15.34	0.04
E69	20.96	0.03
unassigned1 (U1)	24.23	0.05
unassigned2 (U2)	16.24	0.03
Average K_d (μM)		**17.18**
σ K_d (μM)		**4.45**

Experimental Binding Data for T101

Table 52. A table showing the chemical shifts of the T101 residue at increasing concentrations of **157**. The measured and calculated CSP's are also shown.

Lt/Pt	F1 (ppm)	F2 (ppm)	CSP (ppm)	Calculated CSP (ppm)
0	113.18	8.20	0.000	0.000
0.25	113.24	8.21	0.015	0.015
0.5	113.30	8.22	0.031	0.027
1	113.37	8.23	0.045	0.046
2	113.40	8.25	0.058	0.060
4	113.49	8.24	0.065	0.065

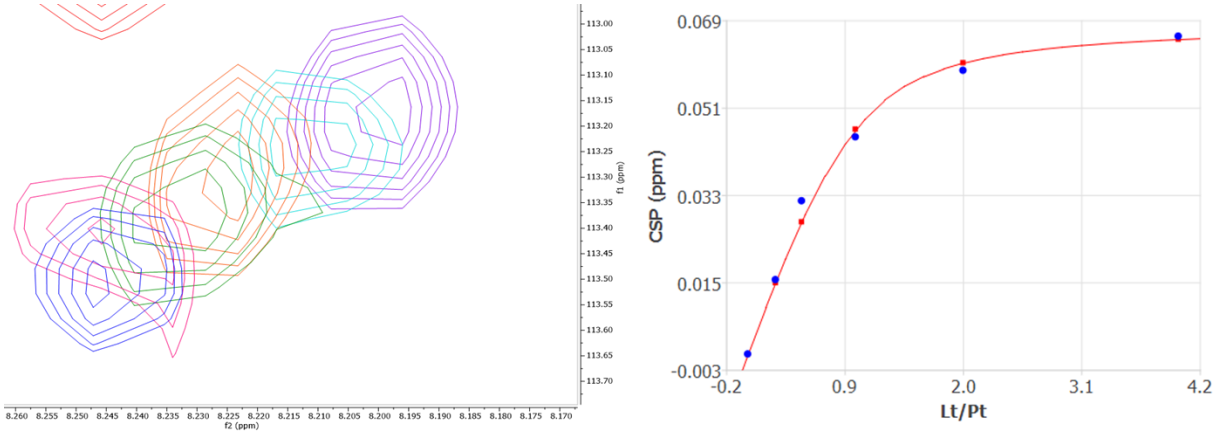


Figure 181. The peaks of the T101 residue at increasing concentrations of **157** (left) and the binding curve plotted using Mbinding. The blue are the experimental CSP's and the red are the calculated CSP's.

Experimental Binding Data for F86

Table 53. A table showing the chemical shifts of the F86 residue at increasing concentrations of **157**. The measured and calculated CSP is also shown.

Lt/Pt	F1 (ppm)	F2 (ppm)	CSP (ppm)	Calculated CSP (ppm)
0	117.01	8.52	0.000	0.000
0.25	117.08	8.51	0.014	0.012
0.5	117.20	8.52	0.030	0.022
1	117.20	8.51	0.030	0.037
2	117.34	8.51	0.053	0.049
4	117.34	8.52	0.052	0.054

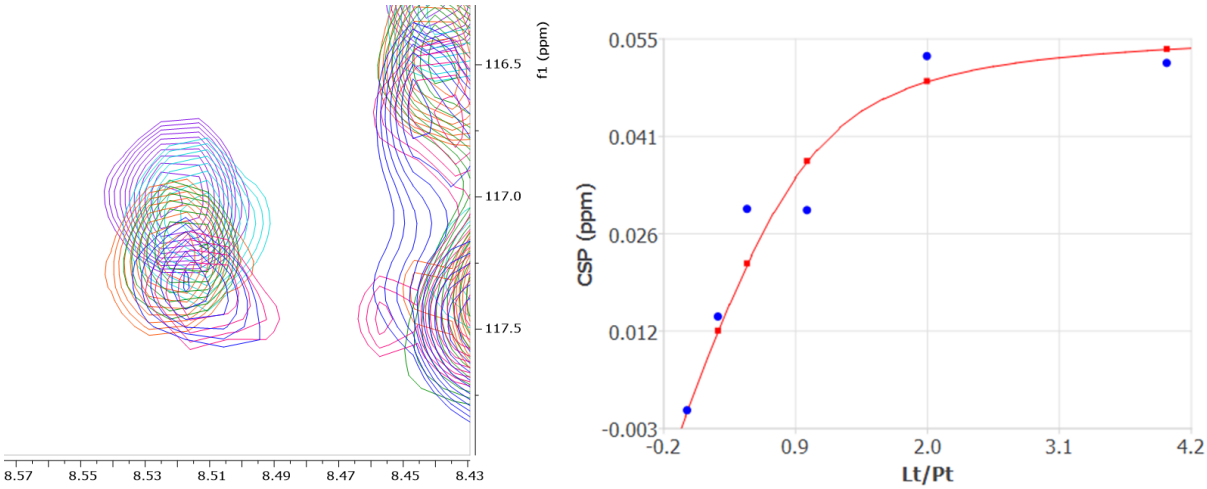


Figure 182. The peaks of the F86 residue at increasing concentrations of **157** (left) and the binding curve plotted using Mbinding. The blue are the experimental CSP's and the red are the calculated CSP's.

Experimental Binding Data for M62

Table 54. A table showing the chemical shifts of the M62 residue at increasing concentrations of **157**. The measured and calculated CSP is also shown.

Lt/Pt	F1 (ppm)	F2 (ppm)	CSP (ppm)	Calculated CSP (ppm)
0	116.17	7.97	0.000	0.000
0.25	116.18	7.98	0.006	0.006
0.5	116.18	7.99	0.012	0.011
1	116.21	7.99	0.018	0.018
2	116.27	7.99	0.025	0.025
4	116.23	8.00	0.029	0.029

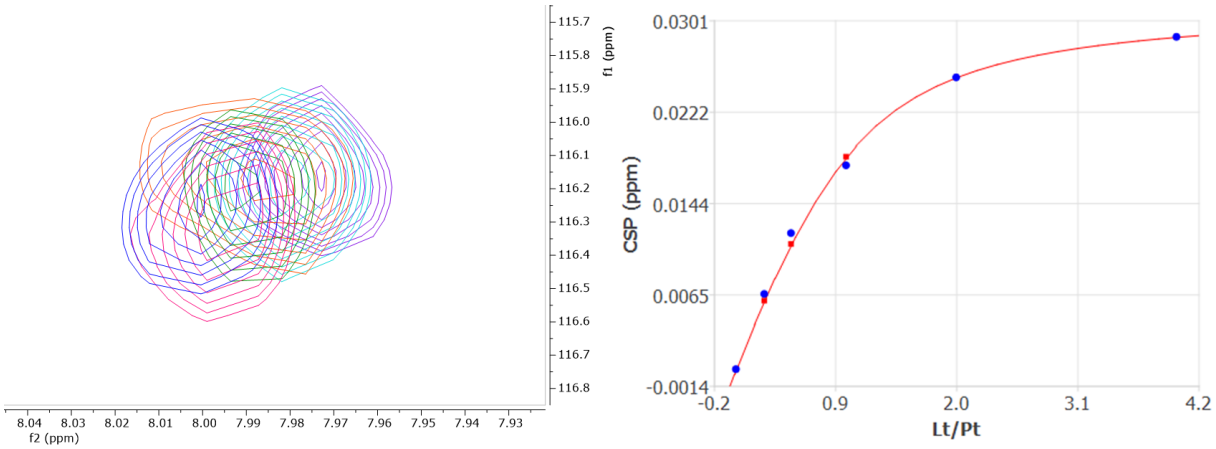
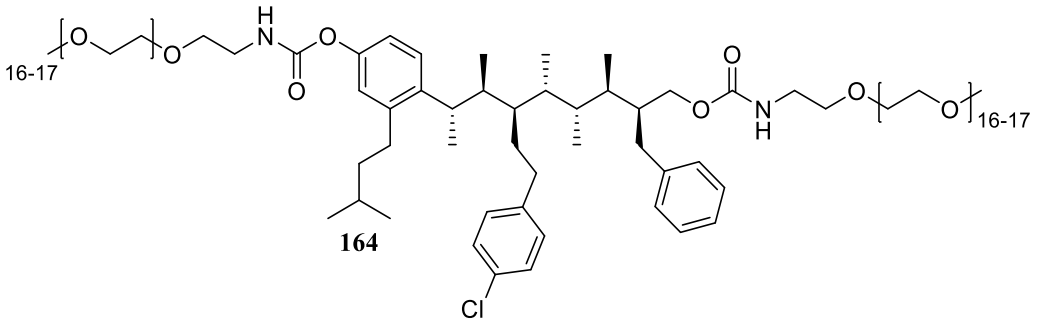


Figure 183. The peaks of the M62 residue at increasing concentrations of **157** (left) and the binding curve plotted using Mbinding. The blue are the experimental CSP's and the red are the calculated CSP's.

9.7.4 The Estimation of K_d using ^{1}H-^{15}N SOFAST HMQC for 164

A series of 2D ^{1}H-^{15}N SOFAST HMQC spectra of Mdm2 (70 μM) with five different concentrations of ligand **164**, ranging from 17.5 to 280 μM, were acquired using the same conditions and procedures as described for ligand **157** previously. **Figure 184** shows an overlay of the spectrum obtained from all five titration points and *apo*-Mdm2. The ^{1}H-^{15}N SOFAST HMQC spectra were analysed using Mbinding from MestreNova and an average value of K_d was obtained across 20 peaks (**Table 55**). Example binding data obtained using Mbinding for three different peaks is also shown below.



164

Table 55. The calculated value of K_d and maximum CSP for 20 different residues interacting with **164**. The average K_d and σ K_d are also shown.

Residue	K_d (μM)	CSP Max (ppm)
C77	11.84	0.09
E69	3.29	0.05
F86	5.43	0.13
F91	7.34	0.06
G83	6.86	0.05
K64	9.70	0.06
K94	6.85	0.11
L107	8.04	0.08
L37	12.41	0.05
L82	11.07	0.04
M62	14.47	0.08
S40	10.62	0.07
S92	3.72	0.13
T47	7.38	0.10
T49	9.96	0.06
U1	7.17	0.07
U2	11.70	0.07
Y104	8.55	0.07
Y48	6.71	0.04
Y60	13.11	0.05
Average K_d (μM)		**8.81**
σ K_d (μM)		**3.07**

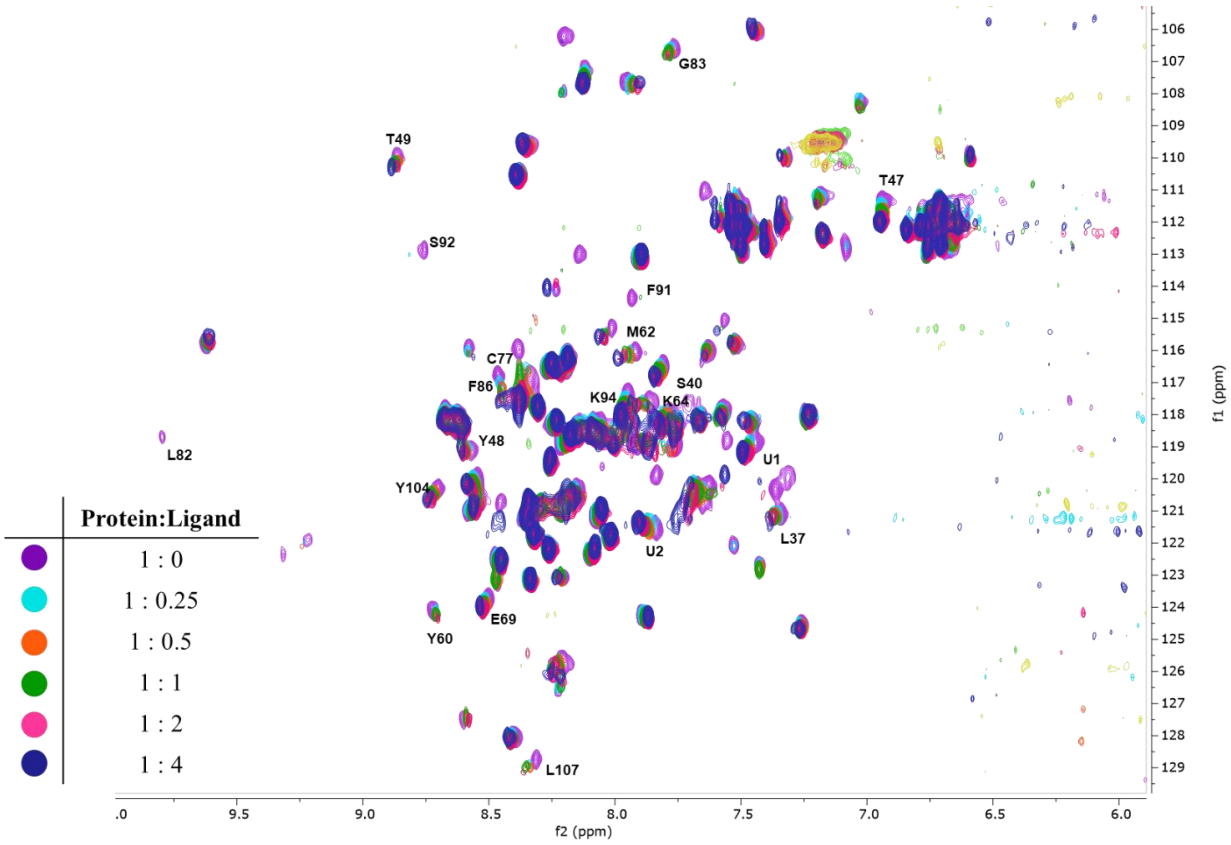


Figure 184. The titration of ligand **164** against Mdm2. Purple (70 μM Mdm2), light blue (70 μM Mdm2 + 17.5 μM ligand), orange (70 μM Mdm2 + 35 μM ligand), green (70 μM Mdm2 + 70 μM ligand), pink (70 μM Mdm2 + 140 μM ligand), dark blue (70 μM Mdm2 + 280 μM ligand). The 20 peaks that were considered in the calculation of K_d are highlighted.

Experimental Binding Data for Y104

Table 56. A table showing the chemical shifts of the Y104 residue at increasing concentrations of **164.** The measured and calculated CSP is also shown.

Lt/Pt	F1 (ppm)	F2 (ppm)	CSP (ppm)	Calculated CSP (ppm)
0	120.32	8.70	0.000	0.000
0.25	120.43	8.72	0.025	0.015
0.5	120.47	8.72	0.030	0.027
1	120.54	8.73	0.043	0.046
2	120.64	8.73	0.058	0.059
4	120.64	8.74	0.063	0.063

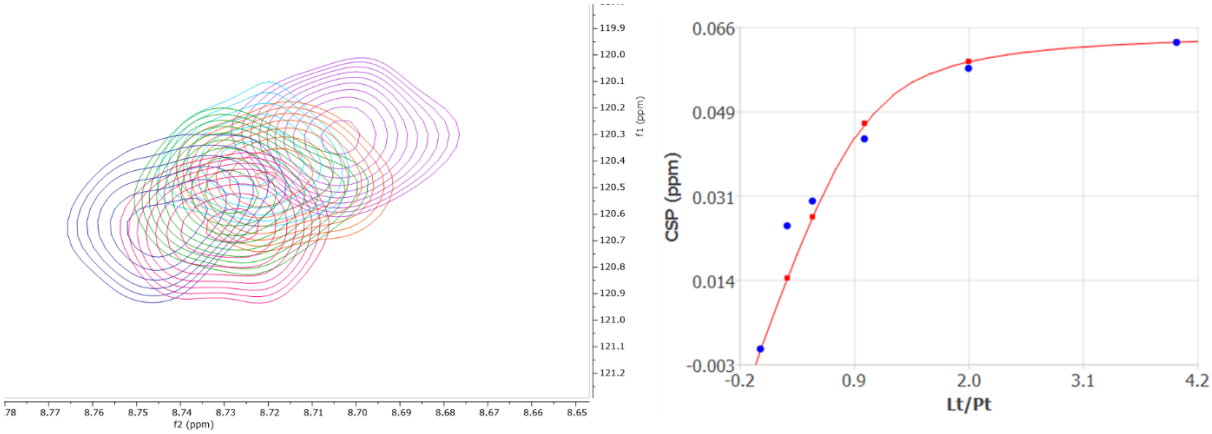


Figure 185. The peaks of the Y104 residue at increasing concentrations of **164** (left) and the binding curve plotted using Mbinding. The blue are the experimental CSP's and the red are the calculated CSP's.

Experimental Binding Data for K94

Table 57. A table showing the chemical shifts of the K94 residue at increasing concentrations of **164**. The measured and calculated CSP is also shown.

Lt/Pt	F1 (ppm)	F2 (ppm)	CSP (ppm)	Calculated CSP (ppm)
0	117.34	7.95	0.000	0.000
0.25	117.63	7.96	0.044	0.025
0.5	117.77	7.96	0.065	0.047
1	117.79	7.97	0.069	0.080
2	117.98	7.97	0.098	0.101
4	118.01	7.98	0.105	0.106

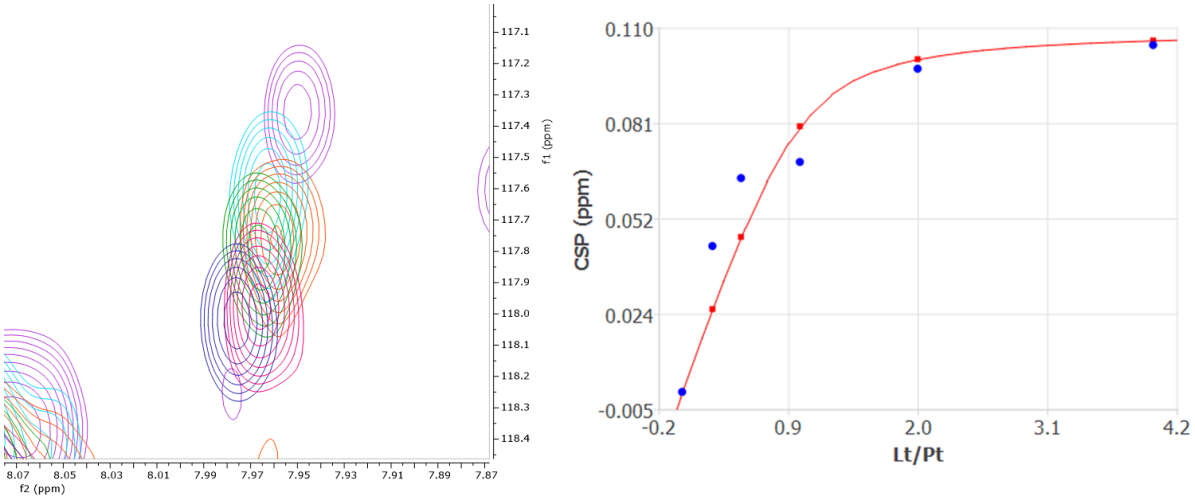


Figure 186. The peaks of the K94 residue at increasing concentrations of **164** (left) and the binding curve plotted using Mbinding. The blue are the experimental CSP's and the red are the calculated CSP's.

Experimental Binding Data for M62

Table 58. A table showing the chemical shifts of the M62 residue at increasing concentrations of **164**. The measured and calculated CSP is also shown.

Lt/Pt	F1 (ppm)	F2 (ppm)	CSP (ppm)	Calculated CSP (ppm)
0	116.03	7.92	0.000	0.000
0.25	116.06	7.94	0.023	0.015
0.5	116.10	7.95	0.031	0.027
1	116.16	7.95	0.043	0.046
2	116.22	7.97	0.061	0.063
4	116.22	7.98	0.072	0.071

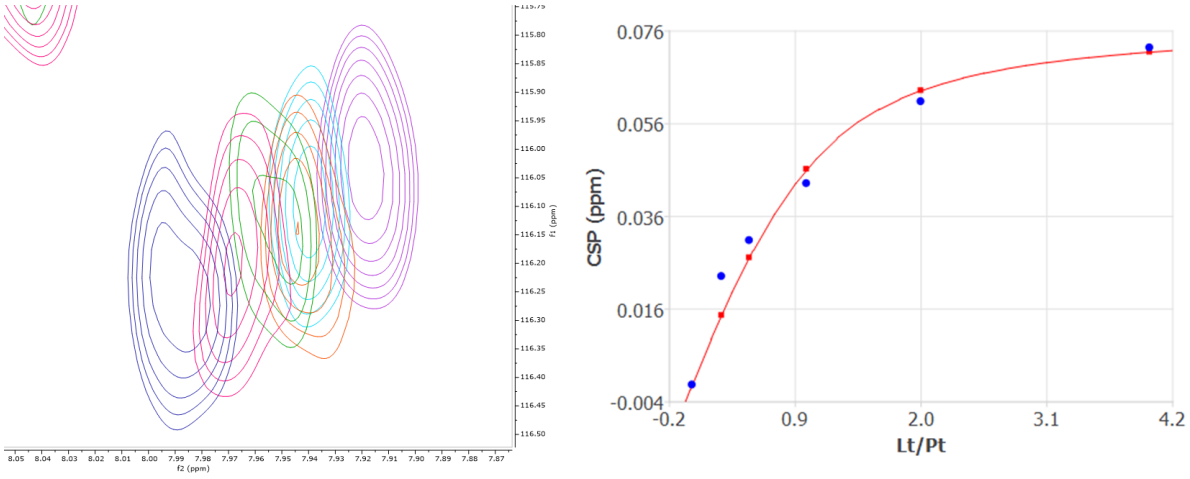
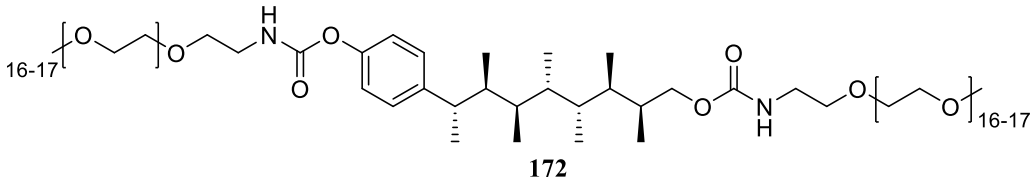


Figure 187. The peaks of the M62 residue at increasing concentrations of **164** (left) and the binding curve plotted using Mbinding. The blue are the experimental CSP's and the red are the calculated CSP's.

9.7.5 The Qualitative Binding of 172 to Mdm2 using ^{1}H-^{15}N SOFAST HMQC

A series of 2D ^{1}H-^{15}N SOFAST HMQC spectra of Mdm2 (70 μM) with five different concentrations of ligand **172**, ranging from 35 to 350 μM, were acquired using the same conditions and procedures as described for ligand **157** previously. **Figure 188** shows an overlay of the spectrum obtained from all five titration points and *apo*-Mdm2. The ^{1}H-^{15}N SOFAST HMQC spectra were analysed using Mbinding from MestreNova. Example binding data obtained using Mbinding for two different peaks is also shown below.



172

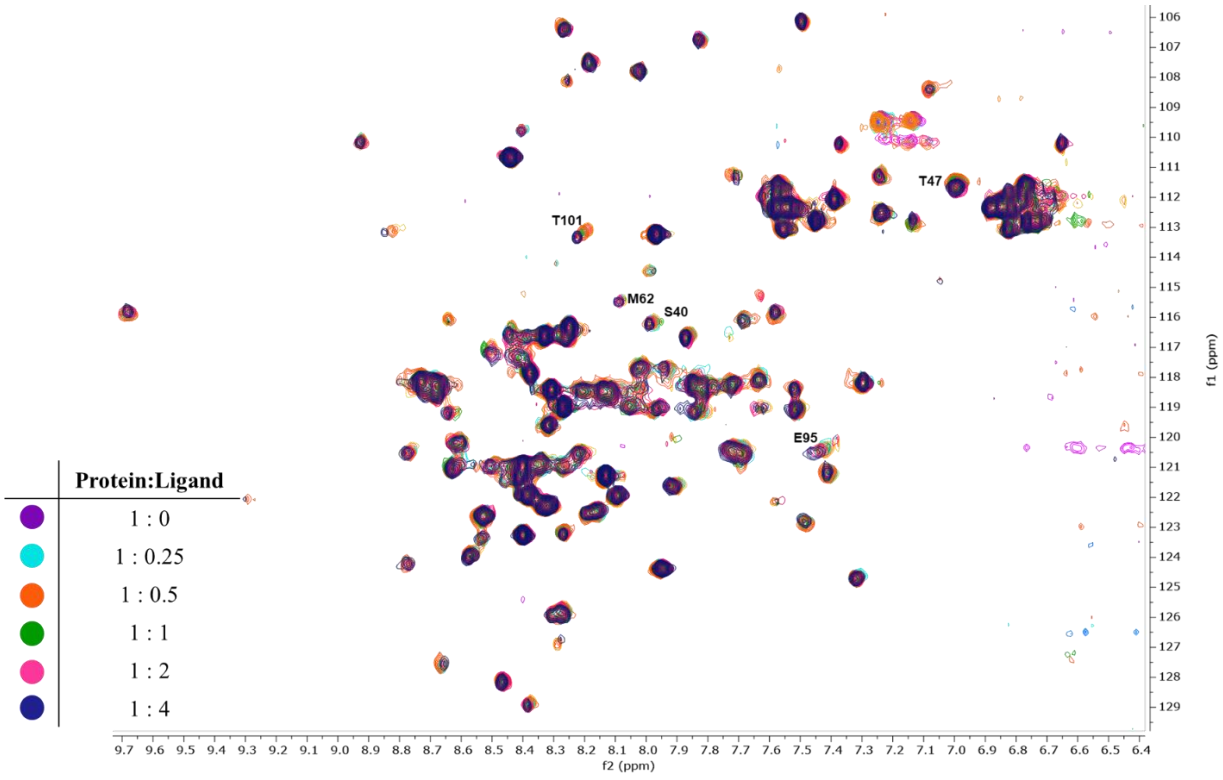


Figure 188. The titration of ligand **172** against Mdm2. Purple (70 μM Mdm2), light blue (70 μM Mdm2 + 35 μM ligand), orange (70 μM Mdm2 + 70 μM ligand), green (70 μM Mdm2 + 140 μM ligand), pink (70 μM Mdm2 + 280 μM ligand), dark blue (70 μM Mdm2 + 350 μM). The 5 peaks that were considered in the calculation of K_d are highlighted.

Example Binding Curve for T101

Table 59. A table showing the chemical shifts of the T101 residue at increasing concentrations of **172.** The measured and calculated CSP is also shown.

Lt/Pt	F1 (ppm)	F2 (ppm)	CSP (ppm)	Calculated CSP (ppm)
0	113.09	8.191	0.000	0.000
0.25	113.17	8.20	0.017	0.011
0.5	113.18	8.21	0.022	0.020
1	113.20	8.21	0.025	0.032
2	113.36	8.22	0.052	0.048
4	113.34	8.23	0.052	0.053

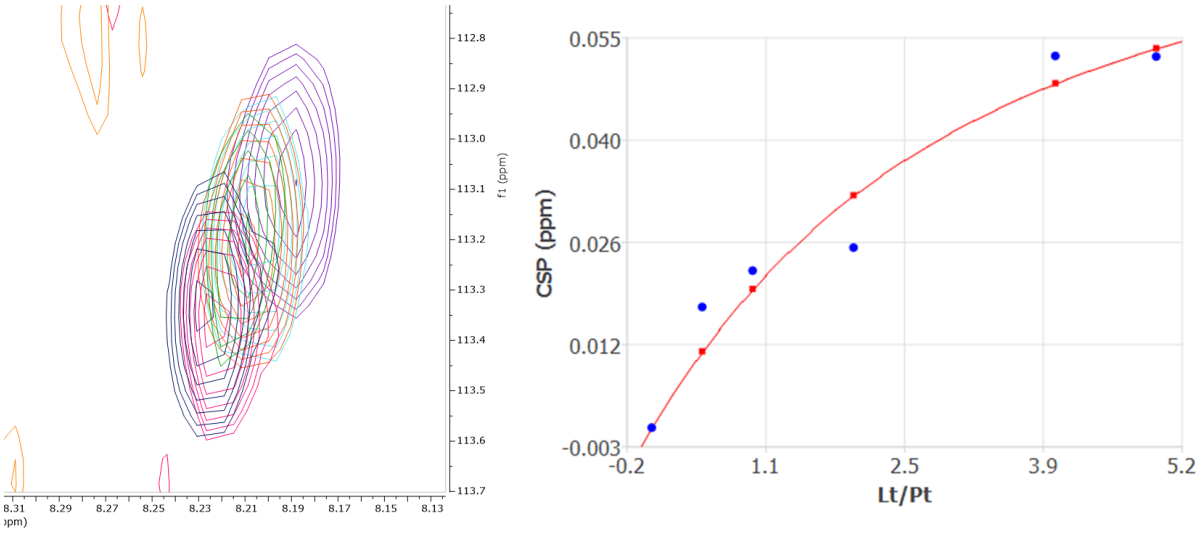


Figure 189. The peaks of the T101 residue at increasing concentrations of **172** (left) and the binding curve plotted using Mbinding. The blue are the experimental CSP's and the red are the calculated CSP's.

Example Binding Curve for M62

Table 60. A table showing the chemical shifts of the M62 residue at increasing concentrations of **172**. The measured and calculated CSP is also shown.

Lt/Pt	F1 (ppm)	F2 (ppm)	CSP (ppm)	Calculated CSP (ppm)
0	116.13	7.98	0.000	0.000
0.25	116.15	7.98	0.004	0.004
0.5	116.18	7.98	0.010	0.008
1	116.14	7.99	0.012	0.013
2	116.22	7.99	0.019	0.020
4	116.24	7.99	0.023	0.022

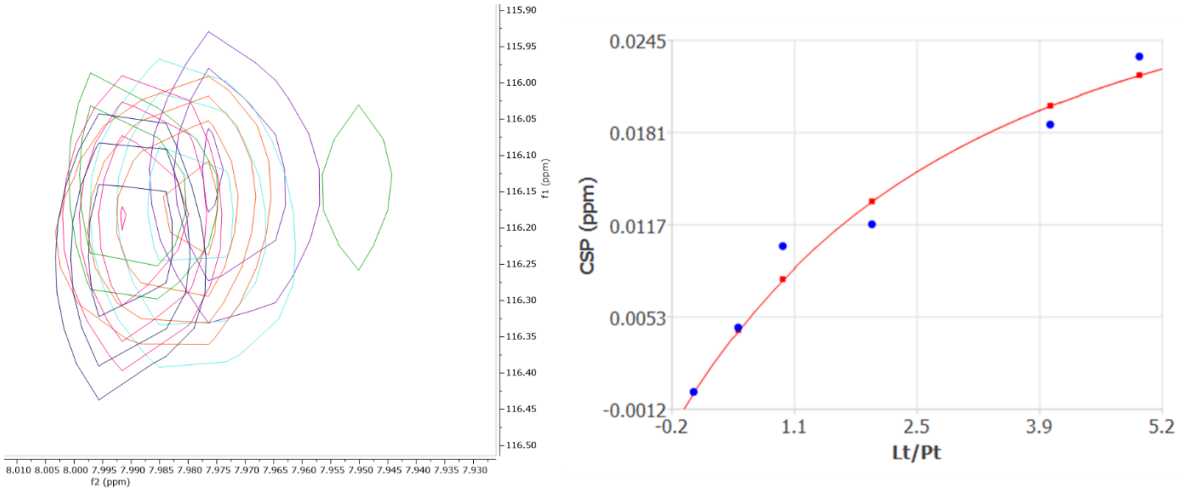


Figure 190. The peaks of the M62 residue at increasing concentrations of **172** (left) and the binding curve plotted using Mbinding. The blue are the experimental CSP's and the red are the calculated CSP's.

9.7.6 The Qualitative Binding of 179 to Mdm2 using ^{1}H-^{15}N SOFAST HMQC

A series of 2D ^{1}H-^{15}N SOFAST HMQC spectra of Mdm2 (70 μM) with six different concentrations of ligand **179**, ranging from 35 to 350 μM, were acquired using the same conditions and procedures as described for ligand **157** previously. **Figure 191** shows an overlay of the spectrum obtained from all six titration points and *apo*-Mdm2. The ^{1}H-^{15}N SOFAST HMQC spectra were analysed using Mbinding from MestreNova. Example binding curves obtained using Mbinding for two different peaks is also shown below.

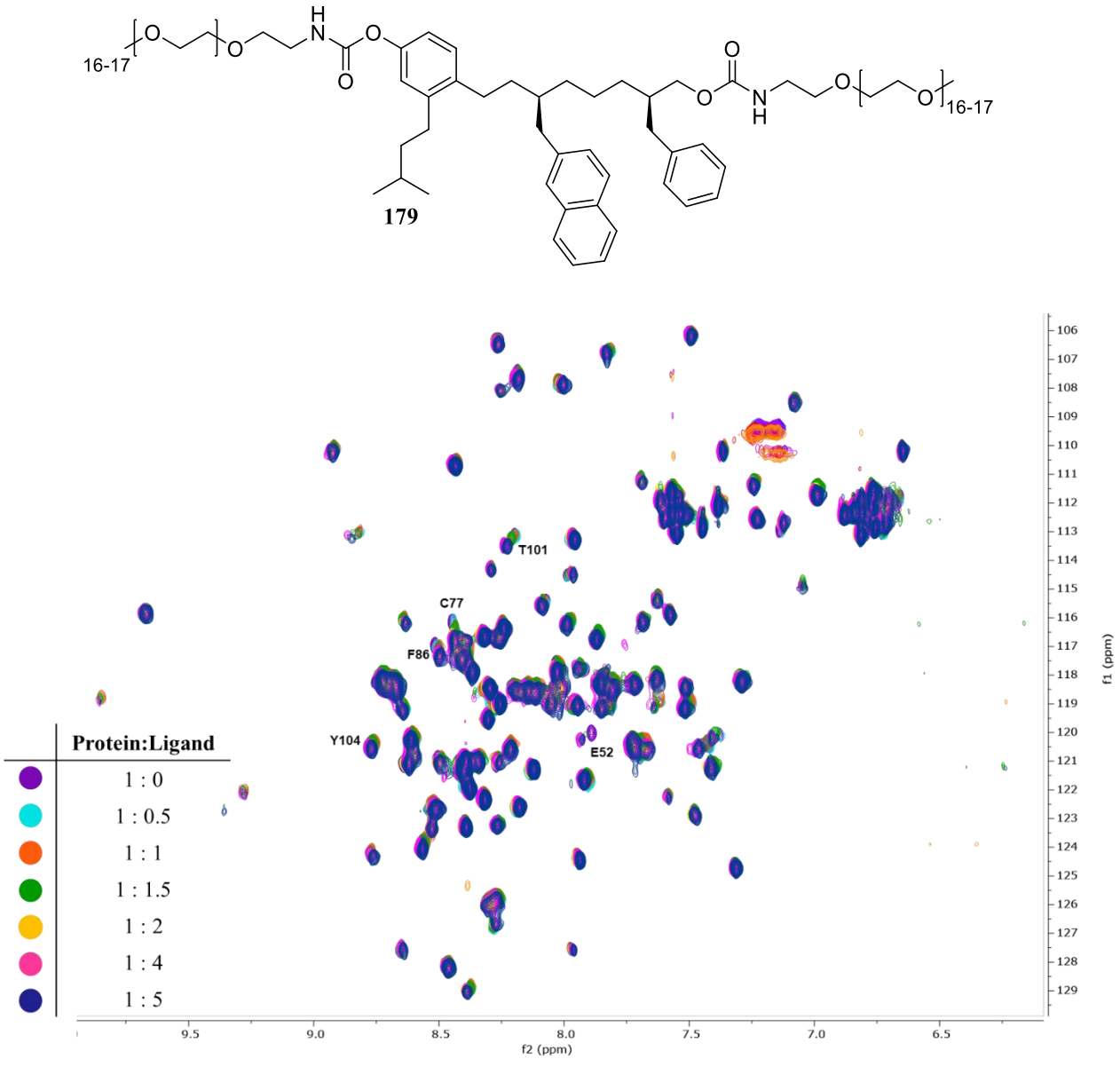


Figure 191. The titration of ligand **179** against Mdm2. Purple (70 μM Mdm2), light blue (70 μM Mdm2 + 35 μM ligand), orange (70 μM Mdm2 + 70 μM ligand), green (70 μM Mdm2 + 105 μM ligand), yellow (70 μM Mdm2 + 140 μM ligand) pink (70 μM Mdm2 + 280 μM ligand), dark blue (70 μM Mdm2 + 350 μM). The 5 peaks that were considered in the calculation of K_d are highlighted.

Example Binding Curve for F86

Table 61. A table showing the chemical shifts of the F86 residue at increasing concentrations of **179**. The measured and calculated CSP is also shown.

Lt/Pt	F1 (ppm)	F2 (ppm)	CSP (ppm)	Calculated CSP (ppm)
0	116.97	8.52	0.000	0.000
0.5	117.01	8.51	0.011	0.012
1	117.05	8.50	0.019	0.022
1.5	117.15	8.50	0.032	0.030
2	117.21	8.51	0.038	0.036
4	117.29	8.50	0.049	0.053
5	117.35	8.50	0.060	0.058

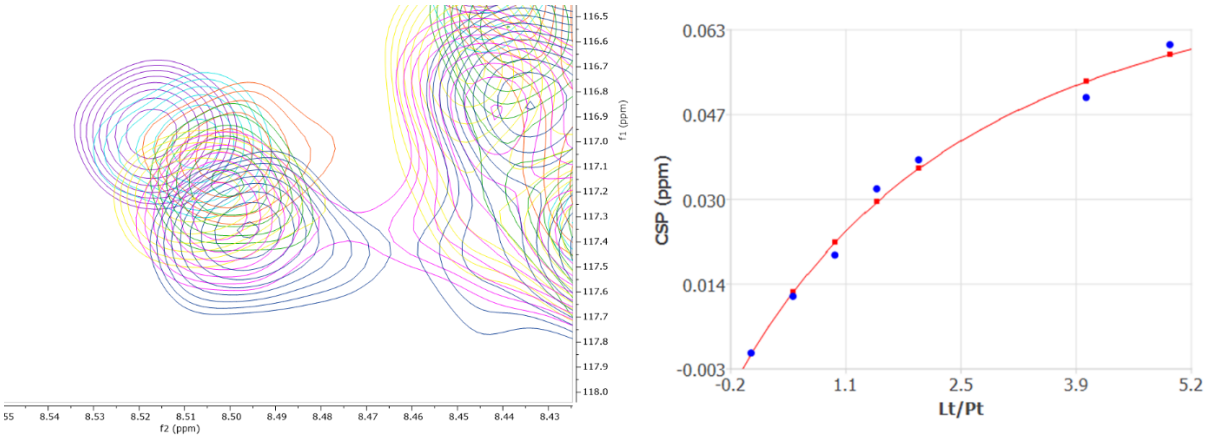


Figure 192. The peaks of the F86 residue at increasing concentrations of **179** (left) and the binding curve plotted using Mbinding. The blue are the experimental CSP's and the red are the calculated CSP's.

Example Binding Curve for C77

Table 62. A table showing the chemical shifts of the F86 residue at increasing concentrations of **179**. The measured and calculated CSP is also shown.

Lt/Pt	F1 (ppm)	F2 (ppm)	CSP (ppm)	Calculated CSP (ppm)
0	116.16	8.45	0.000	0.000
0.5	116.26	8.44	0.016	0.024
1	116.40	8.44	0.036	0.043
1.5	116.55	8.44	0.059	0.058
2	116.69	8.45	0.079	0.071
4	116.88	8.44	0.108	0.102
5	116.86	8.43	0.105	0.112

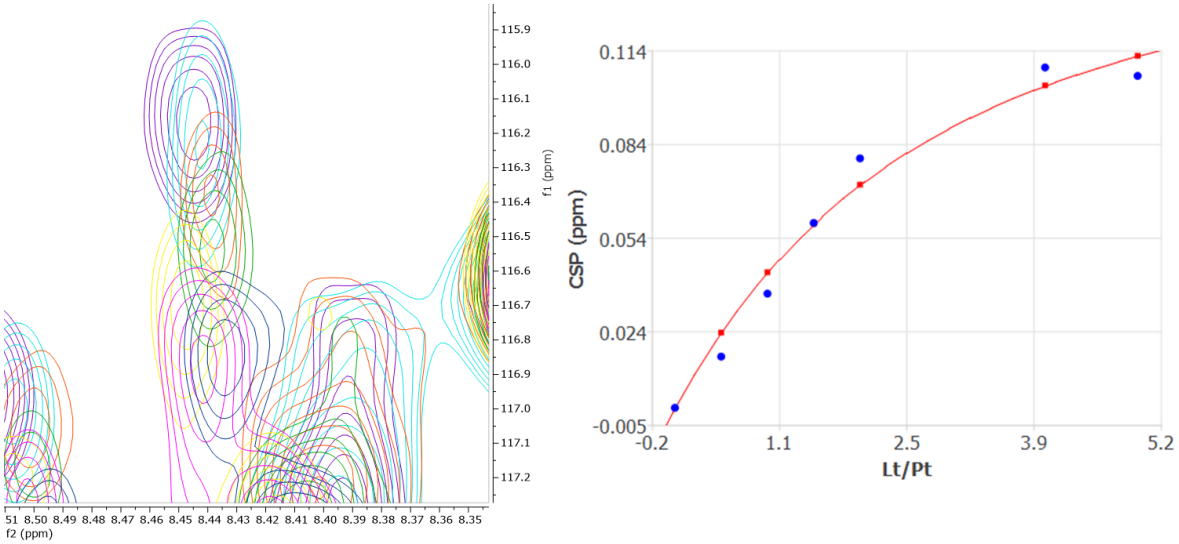


Figure 193. The peaks of the F86 residue at increasing concentrations of **179** (left) and the binding curve plotted using Mbinding. The blue are the experimental CSP's and the red are the calculated CSP's.s

Chapter 10: References

[1] R. W. Hoffmann, *Angew. Chem. Int. Ed.* **2000**, *39*, 2054–2070.

[2] E. Breitmaier, *Structure Elucidation by NMR in Organic Chemistry*, John Wiley & Sons, **2002**.

[3] P. H. Willoughby, M. J. Jansma, T. R. Hoye, *Nat. Protoc.* **2014**, *9*, 643–660.

[4] A. C. Anderson, *Chem. Biol.* **2003**, *10*, 787–797.

[5] R. L. M. Van Montfort, P. Workman, *Essays Biochem.* **2017**, *61*, 431–437.

[6] E. Chiarparin, M. J. Packer, D. M. Wilson, *Future Med. Chem.* **2019**, *11*, 79–82.

[7] V. Lounnas, T. Ritschel, J. Kelder, R. McGuire, R. P. Bywater, N. Foloppe, *Comput. Struct. Biotechnol. J.* **2013**, *5*, 1–14.

[8] Y. Zheng, C. M. Tice, S. B. Singh, *Bioorganic Med. Chem. Lett.* **2017**, *27*, 2825–2837.

[9] E. Perola, P. S. Charifson, *J. Med. Chem.* **2004**, *47*, 2499–2510.

[10] A. E. Tron, M. A. Belmonte, A. Adam, B. M. Aquila, L. H. Boise, E. Chiarparin, J. Cidado, K. J. Embrey, E. Gangl, F. D. Gibbons, et al., *Nat. Commun.* **2018**, *9*, 1–14.

[11] M. Reggelin, H. Hoffmann, M. Köck, D. F. Mierke, *J. Am. Chem. Soc.* **1992**, *114*, 3272–3277.

[12] Z. Fang, Y. Song, P. Zhan, Q. Zhang, X. Liu, *Future Med. Chem.* **2014**, *6*, 885–901.

[13] N. Foloppe, I. J. Chen, *Bioorganic Med. Chem.* **2016**, *24*, 2159–2189.

[14] N. Foloppe, I.-J. Chen, *Future Med. Chem.* **2019**, *11*, 97–118.

[15] E. Perola, P. S. Charifson, *J. Med. Chem* **2004**, *47*, 2499–2510.

[16] S. Michielssens, B. L. De Groot, H. Grubmüller, *Biophys. J.* **2015**, *108*, 2585–2590.

[17] A. M. Aronov, T. Qing, G. Martinez-Botella, G. W. Bemis, J. Cao, G. Chen, N. P. Ewing, P. J. Ford, U. A. Germann, J. Green, et al., *J. Med. Chem.* **2009**, *52*, 6362–6368.

[18] Y. Wang, A. Kirschner, A. K. Fabian, R. Gopalakrishnan, C. Kress, B. Hoogeland, U. Koch, C. Kozany, A. Bracher, F. Hausch, *J. Med. Chem.* **2013**, *56*, 3922–3935.

[19] A. Y. S. Balazs, R. J. Carbajo, N. L. Davies, Y. Dong, A. W. Hird, J. W. Johannes, M. L. Lamb, W. McCoull, P. Raubo, G. R. Robb, et al., *J. Med. Chem.* **2019**, *62*, 9418–9437.

[20] F. Weinhold, *Nature* **2001**, *411*, 539–541.

[21] V. Dragojlovic, *ChemTexts* **2015**, *1*, 1–30.

[22] M. A. Murcko, H. Castejon, K. B. Wiberg, *J. Phys. Chem.* **1996**, *100*, 16162–16168.

[23] G. D. Smith, R. L. Jaffe, *J. Phys. Chem.* **1996**, *100*, 18718–18724.

[24] S. Tsuzuki, L. Schafer, H. Goto, E. D. Jemmis, H. Hosoya, K. Siam, K. Tanabe, E. Osawa, *J. Am. Chem. Soc.* **1991**, 4665–4671.

[25] A. W. Burgess, L. L. Shipman, H. A. Scheraga, *Proc. Natl. Acad. Sci. U. S. A.* **1975**, *72*, 854–858.

[26] R. W. Hoffmann, M. Stahl, U. Schopfer, G. Frenking, *Chem. Eur. J.* **1998**, *4*, 559–566.

[27] M. Burns, S. Essafi, J. R. Bame, S. P. Bull, M. P. Webster, S. Balieu, J. W. Dale, C. P. Butts, J. N. Harvey, V. K. Aggarwal, *Nature* **2014**, *513*, 183–188.

[28] H. Goto, E. Osawa, M. Yamato, *Tetrahedron* **1993**, *49*, 387–396.

[29] J. M. L. Martin, *J. Phys. Chem. A* **2009**, *113*, 11974–11983.

[30] P. W. Smith, W. C. Still, *J. Am. Chem. Soc.* **1988**, *110*, 7917–7919.

[31] J. Mallinson, I. Collins, *Future Med. Chem.* **2012**, *4*, 1409–1438.

[32] C. A. Lipinski, *Adv. Drug Deliv. Rev.* **2016**, *101*, 34–41.

[33] R. W. Hoffmann, *Angew. Chem. Int. Ed.* **1992**, *31*, 1124–1134.

[34] S. K. Malhotra, F. Johnson, *J. Am. Chem. Soc.* **1965**, *87*, 5493–5495.

[35] F. Johnson, *Chem. Rev.* **1968**, *68*, 375–413.

[36] J. L. Broeker, R. W. Hoffmann, K. N. Houk, *J. Am. Chem. Soc.* **1991**, *113*, 5006–5017.

[37] R. W. Hoffmann, *Chem. Rev.* **1989**, *89*, 1841–1860.

[38] R. W. Hoffmann, *Angew. Chem. Int. Ed.* **1992**, *31*, 1124–1134.

[39] R. Göttlich, B. Colin Kahrs, J. Krüger, R. W. Hoffmann, *Chem. Commun.* **1997**, *3*, 247–252.

[40] S. Tsuzuki, L. Schäfer, H. Gotō, E. D. Jemmis, H. Hosoya, K. Siam, K. Tanabe, E. Ōsawa, *J. Am. Chem. Soc.* **1991**, *113*, 4665–4671.

[41] R. H. Boyd, S. M. Breitling, *Macromolecules* **1972**, *5*, 279–286.

[42] D. Leonori, V. K. Aggarwal, *Acc. Chem. Res.* **2014**, *47*, 3174–3183.

[43] J. W. Johannes, S. Bates, C. Beigie, M. A. Belmonte, J. Breen, S. Cao, P. A. Centrella, M. A. Clark, J. W. Cuozzo, C. E. Dumelin, et al., *ACS Med. Chem. Lett.* **2017**, *8*, 239–244.

[44] L. W. Thomas, C. Lam, S. W. Edwards, *FEBS Lett.* **2010**, *584*, 2981–2989.

[45] R. E. Taylor, Y. Chen, G. M. Galvin, P. K. Pabba, *Org. Biomol. Chem.* **2004**, *2*, 127–132.

[46] W. McCoull, R. D. Abrams, E. Anderson, K. Blades, P. Barton, M. Box, J. Burgess, K. Byth, Q. Cao, C. Chuaqui, et al., *J. Med. Chem.* **2017**, *60*, 4386–4402.

[47] E. M. Driggers, S. P. Hale, J. Lee, N. K. Terrett, *Nat. Rev. Drug Discov.* **2008**, *7*, 608–624.

[48] E. Marsault, M. L. Peterson, *J. Med. Chem.* **2011**, *54*, 1961–2004.

[49] D. J. Craik, D. P. Fairlie, S. Liras, D. Price, *Chem. Biol. Drug Des.* **2013**, *81*, 136–147.

[50] J. L. Lau, M. K. Dunn, *Bioorganic Med. Chem.* **2018**, *26*, 2700–2707.

[51] A. M. Ali, J. Atmaj, N. Van Oosterwijk, M. R. Groves, A. Dömling, *Comput. Struct. Biotechnol. J.* **2019**, *17*, 263–281.

[52] L. D. Walensky, G. H. Bird, *J. Med. Chem* **2014**, *57*, 6275−6288.

[53] A. Bhat, L. R. Roberts, J. J. Dwyer, *Eur. J. Med. Chem.* **2015**, *94*, 471–479.

[54] G. L. Verdine, G. J. Hilinski, *Stapled Peptides for Intracellular Drug Targets*, Elsevier Inc., **2012**.

[55] R. Rezaei Araghi, J. A. Ryan, A. Letai, A. E. Keating, *ACS Chem. Biol.* **2016**, *11*, 1238–1244.

[56] M. Llinàs-Brunet, M. Bailey, G. Fazal, S. Goulet, T. Halmos, S. Laplante, R. Maurice, M. Poirier, M. A. Poupart, D. Thibeault, et al., *Bioorganic Med. Chem. Lett.* **1998**, *8*, 1713–1718.

[57] S. R. LaPlante, H. Nar, C. T. Lemke, A. Jakalian, N. Aubry, S. H. Kawai, *J. Med. Chem.* **2014**, *57*, 1777–1789.

[58] K. H. A. Lau, *Biomater. Sci.* **2014**, *2*, 627–633.

[59] R. J. Simon, R. S. Kania, R. N. Zuckermann, V. D. Huebner, D. A. Jewell, S. Banville, S. Ng, L. Wang, S. Rosenberg, C. K. Marlowe, et al., *Proc. Natl. Acad. Sci. U. S. A.* **1992**, *89*, 9367–9371.

[60] G. W. Collie, K. Pulka-Ziach, C. M. Lombardo, J. Fremaux, F. Rosu, M. Decossas, L. Mauran, O. Lambert, V. Gabelica, C. D. Mackereth, et al., *Nat. Chem.* **2015**, *7*, 871–878.

[61] P. S. Farmer, E. J. Ariëns, *Trends Pharmacol. Sci.* **1982**, *3*, 362–365.

[62] B. C. Gorske, E. M. Mumford, C. G. Gerrity, I. Ko, *J. Am. Chem. Soc.* **2017**, *139*, 8070–8073.

[63] O. Roy, C. Caumes, Y. Esvan, C. Didierjean, S. Faure, C. Taillefumier, *Org. Lett.* **2013**, *15*, 2246–2249.

[64] T. Kodadek, P. J. McEnaney, *Chem. Commun.* **2016**, *52*, 6038–6059.

[65] J. Morimoto, Y. Fukuda, D. Kuroda, T. Watanabe, F. Yoshida, M. Asada, T. Nakamura, A. Senoo, S. Nagatoishi, K. Tsumoto, et al., *J. Am. Chem. Soc.* **2019**, *141*, 14612–14623.

[66] A. J. Levine, M. Oren, *Nat. Rev. Cancer* **2009**, *9*, 749–58.

[67] A. D. G. Lawson, M. Maccoss, J. P. Heer, *J. Med. Chem.* **2018**, *61*, 4383–4389.

[68] C. Sandford, V. K. Aggarwal, *Chem. Commun.* **2017**, *53*, 5481–5494.

[69] D. S. Matteson, *J. Organomet. Chem.* **1999**, *581*, 51–65.

[70] V. K. Aggarwal, G. Y. Fang, X. Ginesta, D. M. Howells, M. Zaja, *Pure Appl. Chem.* **2006**, *78*, 215–229.

[71] D. S. Matteson, D. Majumdar, *J. Am. Chem. Soc.* **1980**, *102*, 7588–7590.

[72] D. Majumdar, D. S. Matteson, *J. Am. Chem. Soc.* **1980**, *102*, 7590–7591.

[73] S. P. Thomas, R. M. French, V. Jheengut, V. K. Aggarwal, *Chem. Rec.* **2009**, *9*, 24–39.

[74] S. Essafi, S. Tomasi, V. K. Aggarwal, J. N. Harvey, *J. Org. Chem.* **2014**, *79*, 12148–12158.

[75] J. L. Stymiest, G. Dutheuil, A. Mahmood, V. K. Aggarwal, *Angew. Chem. Int. Ed.* **2007**, *46*, 7491–7494.

[76] E. Beckmann, D. Hoppe, *Synthesis (Stuttg).* **2005**, *2*, 217–222.

[77] E. Beckmann, V. Desai, D. Hoppe, *Synlett* **2004**, *10*, 2275–2280.

[78] R. C. Mykura, S. Veth, A. Varela, L. Dewis, J. J. Farndon, E. L. Myers, V. K. Aggarwal, *J. Am. Chem. Soc.* **2018**, *140*, 14677–14686.

[79] G. Besong, K. Jarowicki, P. J. Kocienski, E. Sliwinski, F. T. Boyle, *Org. Biomol. Chem.* **2006**, *4*, 2193–2207.

[80] P. Beak, M. Baillargeon, L. G. Carter, *J. Org. Chem.* **1978**, *43*, 4255–4256.

[81] P. Beak, L. G. Carter, *J. Org. Chem.* **1981**, *46*, 2363–2373.

[82] R. Larouche-Gauthier, C. J. Fletcher, I. Couto, V. K. Aggarwal, *Chem. Commun.* **2011**, *47*, 12592–4.

[83] M. P. Webster, PhD Thesis. University of Bristol. Multiple Homologations of Boronic Esters and Applications in Synthesis, **2010**.

[84] M. Elyashberg, *TrAC - Trends Anal. Chem.* **2015**, *69*, 88–97.

[85] U. Sternberg, R. Witter, A. S. Ulrich, *Annu. Reports NMR Spectrosc.* **2004**, *52*, 53–104.

[86] G. Bifulco, P. Dambruoso, L. Gomez-Paloma, R. Riccio, *Chem. Rev.* **2007**, *107*, 3744–3779.

[87] M. Karplus, *J. Chem. Phys.* **1959**, *30*, 11–15.

[88] M. Karplus, *J. Am. Chem. Soc.* **1963**, *85*, 2870–2871.

[89] C. A. G. Haasnoot, F. A. A. M. de Leeuw, C. Altona, *Tetrahedron* **1980**, *36*, 2783–2792.

[90] C. A. G. Haasnoot, F. A. A. M. De Leeuw, H. P. M. De Leeuw, C. Altona, *Biopolymers* **1981**, *20*, 1211–1245.

[91] C. Altona, J. H. Ippel, A. J. A. W. Hoekzema, C. Erkelens, M. Groesbeek, L. A. Donders, *Magn. Reson. Chem.* **1989**, *27*, 564–576.

[92] G. J. Karabatsos, C. E. Orzech, N. Hsi, *J. Am. Chem. Soc.* **1966**, *88*, 1817–1818.

[93] M. J. Minch, *Concepts Magn. Reson. Part A* **1994**, *6*, 41–56.

[94] M. Foroozandeh, R. W. Adams, P. Kiraly, M. Nilsson, G. A. Morris, *Chem. Commun.* **2015**, *51*, 15410–15413.

[95] P. Kiraly, M. Foroozandeh, M. Nilsson, G. A. Morris, *Chem. Phys. Lett.* **2017**, *683*, 398–403.

[96] N. Matsumori, D. Kaneno, M. Murata, H. Nakamura, K. Tachibana, *J. Org. Chem.* **1999**, *64*, 866–876.

[97] J. L. Marshall, D. E. Müller, S. A. Conn, R. Seiwell, *Acc. Chem. Res.* **1974**, *7*, 333–339.

[98] T. Parella, J. F. Espinosa, *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 17–55.

[99] C. L. Dickson, C. D. Blundell, C. P. Butts, A. Felton, A. Jeffreys, Z. Takacs, *Analyst* **2017**, *142*, 621–633.

[100] J. Saurí, T. Parella, *Magn. Reson. Chem.* **2013**, *51*, 509–516.

[101] S. Gil, J. F. Espinosa, T. Parella, *J. Magn. Reson.* **2010**, *207*, 312–321.

[102] C. P. Butts, C. R. Jones, J. N. Harvey, *Chem. Commun.* **2011**, *47*, 1193–1195.

[103] L. Dewis, R. Crouch, D. Russell, C. Butts, *Magn Reson Chem* **2019**, *57*, 1143–1149.

[104] C. P. Butts, C. R. Jones, E. C. Towers, J. L. Flynn, L. Appleby, N. J. Barron, *Org. Biomol. Chem.* **2011**, *9*, 177–184.

[105] G. Bodenhausen, R. R. Ernst, *J. Am. Chem. Soc* **1982**, *104*, 1304–1309.

[106] S. Macura, B. T. Farmer, L. R. Brown, *J. Magn. Reson.* **1986**, *70*, 493–499.

[107] H. Hu, K. Krishnamurthy, *J. Magn. Reson.* **2006**, *182*, 173–177.

[108] M. G. Chini, C. R. Jones, A. Zampella, M. V. D'Auria, B. Renga, S. Fiorucci, C. P. Butts, G. Bifulco, *J. Org. Chem.* **2012**, *77*, 1489–1496.

[109] S. Di Micco, A. Zampella, M. V. D'Auria, C. Festa, S. De Marino, R. Riccio, C. P. Butts, G. Bifulco, *Beilstein J. Org. Chem.* **2013**, *9*, 2940–2949.

[110] G. Barone, L. Gomez-Paloma, D. Duca, A. Silvestri, R. Riccio, G. Bifulco, *Chem. Eur. J.* **2002**, *8*, 3233–3239.

[111] I. Chen, N. Foloppe, *Drug Dev. Res.* **2011**, *94*, 85–94.

[112] T. Kamachi, K. Yoshizawa, *J. Chem. Inf. Model* **2016**, *56*, 347–353.

[113] J. Schlitter, *J. Mol. Graph.* **1994**, *12*, 84–89.

[114] D. K. Agrafiotis, A. C. Gibbs, F. Zhu, S. Izrailev, E. Martin, *J. Chem. Inf. Model* **2007**, *47*, 1067–1086.

[115] R. G. Woolley, *chem phys lett* **1977**, *45*, 393–398.

[116] M. Molecular, F. Field, T. A. Halgren, *J. Comput. Chem.* **1996**, *17*, 490–519.

[117] V. T. Terms, G. R. March, N. L. Allinger, *J. Am. Chem. Soc.* **1977**, *99*, 8127–8134.

[118] L. Norman, H. Young, *J. Am. Chem. Soc.* **1989**, *111*, 8551–8566.

[119] G. A. Kaminski, R. A. Friesner, J. Tirado-rives, W. L. Jorgensen, *J. Phys. Chem. B* **2001**, *2*, 6474–6487.

[120] G. Chang, W. C. Guida, W. C. Still, *J. Am. Chem. Soc.* **1989**, *111*, 4379–4386.

[121] D. Bagayoko, *AIP Adv.* **2014**, *4*, 1–12.

[122] A. D. Becke, *J. Chem. Phys.* **2014**, *140*, 0–18.

[123] A. D. Becke, *J. Chem. Phys.* **1993**, *98*, 5648–5652.

[124] P. J. Stephens, F. J. Devlin, C. F. Chabalowski, M. J. Frisch, *J. Phys. Chem.* **1994**, *98*, 11623–11627.

[125] C. Adamo, V. Barone, *J. Chem. Phys* **1998**, *108*, 664–675.

[126] K. Burke, *J. Chem. Phys.* **2012**, *136*, 1–9.

[127] C. J. Cramer, *Essentials of Computational Chemistry*, **2004**.

[128] R. Ditchfield, W. J. Hehre, J. A. Pople, *J. Chem. Phys.* **1971**, *54*, 724–728.

[129] J. A. Montgomery, M. J. Frisch, J. W. Ochterski, G. A. Petersson, *J. Chem. Phys.* **1999**, *110*, 2822–2827.

[130] M. J. Frisch, J. A. Pople, J. S. Binkley, *J. Chem. Phys.* **1984**, *80*, 3265–3269.

[131] A. Bagno, *Chem. - A Eur. J.* **2001**, *7*, 1652–1661.

[132] A. Bagno, F. Rastrelli, G. Saielli, *Chem. - A Eur. J.* **2006**, *12*, 5514–5525.

[133] N. Grimblat, A. M. Sarotti, *Chem. - A Eur. J.* **2016**, *22*, 12246–12261.

[134] S. K. Wolff, T. Ziegler, *J. Chem. Phys.* **1998**, *109*, 895–905.

[135] S. G. Smith, J. M. Goodman, *J. Am. Chem. Soc.* **2010**, *132*, 12946–12959.

[136] S. Di Micco, M. G. Chini, R. Riccio, G. Bifulco, *European J. Org. Chem.* **2010**, *8*, 1411–1434.

[137] G. Barone, D. Duca, A. Silvestri, L. Gomez-Paloma, R. Riccio, G. Bifulco, *Chem. - A Eur. J.* **2002**, *8*, 3240–3245.

[138] P. Cimino, L. Gomez-Paloma, D. Duca, R. Riccio, G. Bifulco, *Magn. Reson. Chem.* **2004**, *42*, 26–33.

[139] T. A. Palazzo, T. T. Truong, S. M. T. Wong, E. T. Mack, M. W. Lodewyk, J. G. Harrison, R. A. Gamage, J. B. Siegel, M. J. Kurth, D. J. Tantillo, *J. Chem. Educ.* **2015**, *92*, 561–566.

[140] S. Zaretsky, J. L. Hickey, M. A. St, C. C. G. Scully, A. L. Roughton, D. J. Tantillo, M. W. Lodewyk, A. K. Yudin, *Tetrahedron* **2014**, *70*, 7655–7663.

[141] Q. N. N. Nguyen, J. Schwochert, D. J. Tantillo, R. S. Lokey, *Phys. Chem. Chem. Phys.* **2018**, *20*, 14003–14012.

[142] K. Ermanis, K. E. B. Parkes, J. M. Goodman, T. Agback, *Org. Biomol. Chem.* **2019**, *17*, 5886–5890.

[143] J. De Las Rivas, C. Fontanillo, M. Furumichi, M. Tanabe, M. Hirakawa, *PLoS Comput. Biol.* **2010**, *6*, 1–8.

[144] I. M. A. Nooren, J. M. Thornton, *Eur. Mol. Biol. Organ.* **2003**, *22*, 3486–3492.

[145] V. Azzarito, K. Long, N. S. Murphy, A. J. Wilson, *Nat. Chem.* **2013**, *5*, 161–173.

[146] H. Yin, A. D. Hamilton, *Angew. Chem. Int. Ed.* **2005**, *44*, 4130–4163.

[147] C. V. Dang, E. P. Reddy, K. M. Shokat, L. Soucek, *Nat. Rev. Cancer* **2017**, *17*, 502–508.

[148] X. Ran, J. E. Gestwicki, *Curr. Opin. Chem. Biol.* **2018**, *44*, 75–86.

[149] L. Mabonga, A. P. Kappo, *Biophys. Rev.* **2019**, *11*, 559–581.

[150] M. Bakail, F. Ochsenbein, *Comptes Rendus Chim.* **2016**, *19*, 19–27.

[151] A. Groß, C. Hashimoto, H. Sticht, J. Eichler, *Front. Bioeng. Biotechnol.* **2016**, *3*, 1–16.

[152] S. Lien, H. B. Lowman, *Trends Biotechnol.* **2003**, *21*, 556–562.

[153] K. Fosgerau, T. Hoffmann, *Drug Discov. Today* **2015**, *20*, 122–128.

[154] T. Clackson, J. A. Wells, *Science.* **1995**, *267*, 383–386.

[155] J. Kenneth Morrow, S. Zhang, *Curr. Drug Metab.* **2012**, *18*, 1255–1265.

[156] E. Cukuroglu, H. B. Engin, A. Gursoy, O. Keskin, *Prog. Biophys. Mol. Biol.* **2014**, *116*, 165–173.

[157] M. Guharoy, P. Chakrabarti, *BMC Bioinformatics* **2010**, *11*, 1–17.

[158] M. R. Arkin, Y. Tang, J. A. Wells, *Chem. Biol.* **2014**, *21*, 1102–1114.

[159] A. Mullard, *Nat. Rev. Drug Discov.* **2012**, *11*, 173–175.

[160] A. J. Wilson, *Chem. Soc. Rev.* **2009**, *38*, 3289–3300.

[161] D. P. Lane, S. Benchimol, *Genes Dev.* **1990**, *4*, 1–8.

[162] D. P. Lane, *Nature* **1992**, *358*, 15–16.

[163] J. Bargonetti, C. Prives, *J. Mol. Cell Biol.* **2019**, *11*, 605–609.

[164] B. Vogelstein, D. Lane, A. J. Levine, *Nature* **2000**, *408*, 307–310.

[165] N. Baptiste, P. Friedlander, X. Chen, C. Prives, *Oncogene* **2002**, *21*, 9–21.

[166] J. Zhu, W. Zhou, J. Jiang, X. Chen, *J. Biol. Chem.* **1998**, *273*, 13030–13036.

[167] S. E. Kern, K. W. Kinzler, A. Bruskin, D. Jarosz, P. Friedman, C. Prives, B. Vogelstein, *Science.* **1991**, *252*, 1708–1711.

[168] K. G. McLure, P. W. K. Lee, *EMBO J.* **1998**, *17*, 3342–3350.

[169] P. Chène, *Oncogene* **2001**, *20*, 2611–2617.

[170] W. Gu, R. G. Roeder, *Cell* **1997**, *90*, 595–606.

[171] A. M. Goh, C. R. Coffill, D. P. Lane, *J. Pathol.* **2011**, *223*, 116–126.

[172] S. W. Chi, S. H. Lee, D. H. Kim, M. J. Ahn, J. S. Kim, J. Y. Woo, T. Torizawa, M. Kainosho, K. H. Han, *J. Biol. Chem.* **2005**, *280*, 38795–38802.

[173] O. Schon, A. Friedler, M. Bycroft, S. M. V Freund, A. R. Fersht, *J. Mol. Biol.* **2002**, *323*, 491–501.

[174] H. Lee, K. H. Mok, R. Muhandiram, K. H. Park, J. E. Suk, D. H. Kim, J. Chang, Y. C. Sung, K. Y. Choi, K. H. Han, *J. Biol. Chem.* **2000**, *275*, 29426–29432.

[175] X. Wu, J. H. Bayle, D. Olson, A. J. Levine, *Genes Dev.* **1993**, *53*, 1126–1132.

[176] J. Momand, H.-H. Wu, G. Dasgupta, *Gene* **2000**, *242*, 15–29.

[177] Y. Haupt, R. Maya, A. Kazaz, M. Oren, *Nature* **1997**, *387*, 296–9.

[178] M. H. G. Kubbutat, S. N. Jones, K. H. Vousden, *Nature* **1997**, *387*, 299–303.

[179] U. M. Moll, O. Petrenko, *Mol. Cancer Res.* **2003**, *1*, 1001–1008.

[180] S. Y. Shieh, J. Ahn, K. Tamai, Y. Taya, C. Prives, *Genes Dev.* **2000**, *14*, 289–300.

[181] Y. Liu, O. Tavana, W. Gu, *J. Mol. Cell Biol.* **2019**, *11*, 564–577.

[182] S. L. Harris, A. J. Levine, *Oncogene* **2005**, *24*, 2899–2908.

[183] P. A. J. Muller, K. H. Vousden, *Nat. Cell Biol.* **2013**, *15*, 2–8.

[184] M. Hollstein, D. Sidransky, B. Vogelstein, C. C. Harris, *Science.* **1991**, *253*, 49–54.

[185] A. Sigal, V. Rotter, *Cancer Res.* **2000**, *60*, 6788–6793.

[186] N. Rivlin, R. Brosh, M. Oren, V. Rotter, *Genes and Cancer* **2011**, *2*, 466–474.

[187] Ö. Demir, P. U. Ieong, R. E. Amaro, *Oncogene* **2017**, *36*, 1451–1460.

[188] A. Perdrix, A. Najem, S. Saussez, A. Awada, F. Journe, G. Ghanem, M. Krayem, *Cancers (Basel).* **2017**, *9*, 1–17.

[189] V. J. N. Bykov, N. Issaeva, A. Shilov, M. Hultcrantz, E. Pugacheva, P. Chumakov, J. Bergman, K. G. Wiman, G. Selivanova, *Nat. Med.* **2002**, *8*, 282–288.

[190] J. M. R. Lambert, P. Gorzov, D. B. Veprintsev, M. Söderqvist, D. Segerbäck, J. Bergman, A. R. Fersht, P. Hainaut, K. G. Wiman, V. J. N. Bykov, *Cancer Cell* **2009**, *15*, 376–388.

[191] C. D. Wassman, R. Baronio, Ö. Demir, B. D. Wallentine, C.-K. Chen, L. V Hall, F. Salehi, D.-W. Lin, B. P. Chung, G. W. Hatfield, et al., *Nat. Commun.* **2013**, *4*, 1407.

[192] S. N. Jones, A. R. Hancock, H. Vogel, L. A. Donehower, A. Bradley, *Proc. Natl. Acad. Sci. U. S. A.* **1998**, *95*, 15608–15612.

[193] P. M. Fischer, D. P. Lane, *Trends Pharmacol. Sci.* **2004**, *25*, 343–346.

[194] P. H. Kussie, S. Gorina, V. Marechal, B. Elenbaas, J. Moreau, A. J. Levine, N. P. Pavletich, *a.* **1996**, *274*, 948–953.

[195] L. T. Vassilev, B. T. Vu, B. Craves, D. Carvajal, F. Podlaski, Z. Filipovic, N. Kong, U. Kammlott, C. Lukacs, C. Klein, et al., *Science.* **2004**, *303*, 844–848.

[196] Y. Zhao, A. Aguilar, D. Bernard, S. Wang, *J. Med. Chem.* **2015**, *58*, 1038–1052.

[197] B. Vu, P. Wovkulich, G. Pizzolato, A. Lovey, Q. Ding, N. Jiang, J. J. Liu, C. Zhao, K. Glenn, Y. Wen, et al., *Med. Chem. Lett.* **2013**, *4*, 466–469.

[198] A. Michael, K. R. Kelly, K. Yee, S. Assouline, R. Strair, L. Popplewell, D. Bowen, G. Martinelli, M. W. Drummond, P. Vyas, et al., *Clin Cancer Res* **2016**, *22*, 868–876.

[199] S. Shangary, D. Qin, D. McEachern, M. Liu, R. S. Miller, S. Qiu, Z. Nikolovska-Coleska, K. Ding, G. Wang, J. Chen, et al., *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 3933–8.

[200] Q. Ding, Z. Zhang, J.-J. Liu, N. Jiang, J. Zhang, T. M. Ross, X.-J. Chu, D. Bartkovitz, F. Podlaski, C. Janson, et al., *J. Med. Chem.* **2013**, *56*, 5979–5983.

[201] S. Wang, Y. Zhao, A. Aguilar, D. Bernard, C. Y. Yang, *Cold Spring Harb. Perspect. Med.* **2017**, *7*, 1–11.

[202] J. Lu, D. McEachern, S. Li, M. J. Ellis, S. Wang, *Mol. Cancer Ther.* **2016**, *15*, 2887–2893.

[203] H. Yin, G. I. Lee, S. P. Hyung, G. A. Payne, J. M. Rodriguez, S. M. Sebti, A. D. Hamilton, *Angew. Chemie - Int. Ed.* **2005**, *44*, 2704–2707.

[204] M. K. P. Jayatunga, S. Thompson, A. D. Hamilton, *Bioorganic Med. Chem. Lett.* **2014**, *24*, 717–724.

[205] B. N. Bullock, A. L. Jochim, P. S. Arora, *J. Am. Chem. Soc.* **2011**, *133*, 14220–14223.

[206] A. L. Jochim, P. S. Arora, *ACS Chem. Biol.* **2010**, *5*, 919–923.

[207] L. K. Henchey, A. L. Jochim, P. S. Arora, *Curr. Opin. Chem. Biol.* **2008**, *12*, 692–697.

[208] M. Raj, B. N. Bullock, P. S. Arora, *Bioorganic Med. Chem.* **2013**, *21*, 4051–4057.

[209] B. P. Orner, J. T. Ernst, A. D. Hamilton, *J. Am. Chem. Soc.* **2001**, *123*, 5382–5383.

[210] J. T. Ernst, O. Kutzki, A. K. Debnath, S. Jiang, H. Lu, A. D. Hamilton, *Angew. Chem. Int. Ed.* **2002**, *41*, 278–281.

[211] H. Yin, G. I. Lee, K. A. Sedey, O. Kutzki, H. S. Park, B. P. Orner, J. T. Ernst, H. G. Wang, S. M. Sebti, A. D. Hamilton, *J. Am. Chem. Soc.* **2005**, *127*, 10191–10196.

[212] C. G. Cummings, A. D. Hamilton, *Curr. Opin. Chem. Biol.* **2010**, *14*, 341–346.

[213] O. Kutzki, H. S. Park, J. T. Ernst, B. P. Orner, H. Yin, A. D. Hamilton, *J. Am. Chem. Soc.* **2002**, *124*, 11838–11839.

[214] J. M. Davis, A. Truong, A. D. Hamilton, *Org. Lett.* **2005**, *7*, 5405–5408.

[215] H. Yin, G. Lee, K. A. Sedey, J. M. Rodriguez, H. Wang, S. M. Sebti, A. D. Hamilton, Y. U. V, P. O. Box, *J. Am. Chem. Soc.* **2005**, *127*, 5463–5468.

[216] D. Vodak, J. Wachter, M. O. Keeffe, O. M. Yaghi, P. J. Hagraman, J. Zubieta, A. Chem, M. O. Keeffe, O. M. Yaghi, J. A. C. Soc, et al., *Angew. Chem. Int. Ed.* **2003**, *42*, 535–539.

[217] I. C. Kim, A. D. Hamilton, *Org. Lett.* **2006**, *8*, 1751–1754.

[218] C. G. Cummings, N. T. Ross, W. P. Katt, A. D. Hamilton, *Org. Lett.* **2009**, *11*, 25–28.

[219] J. M. Rodriguez, L. Nevola, N. T. Ross, G. Lee, A. D. Hamilton, *Chem. Bio. Chem* **2009**, *10*, 829–833.

[220] J. M. Rodriguez, A. D. Hamilton, *Tetrahedron Lett.* **2006**, *47*, 7443–7446.

[221] J. M. Rodriguez, A. D. Hamilton, *Angew. Chem. Int. Ed.* **2007**, *46*, 8614–8617.

[222] J. M. Davis, L. K. Tsou, A. D. Hamilton, *Chem. Soc. Rev.* **2007**, *36*, 326–334.

[223] S. Fletcher, A. D. Hamilton, *J. R. Soc. Interface* **2006**, *3*, 215–233.

[224] Y. Lee, H. Im, S. Das, M. Oh, J. H. Lee, S. Ham, H. S. Lim, *Chem. Commun.* **2019**, *55*, 13311–13314.

[225] M. Oh, J. H. Lee, W. Wang, H. S. Lee, W. S. Lee, C. Burlak, W. Im, Q. Q. Hoang, H. S. Lim, *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 11007–11012.

[226] H. Oguri, A. Oomura, S. Tanabe, M. Hirama, *Tetrahedron Lett.* **2005**, *46*, 2179–2183.

[227] S. Thompson, R. Vallinayagam, M. J. Adler, R. T. W. Scott, A. D. Hamilton, *Tetrahedron* **2012**, *68*, 4501–4505.

[228] S. Thompson, A. D. Hamilton, *Org. Biomol. Chem.* **2012**, *10*, 5780–5782.

[229] K. Jung, K. Vanommeslaeghe, M. E. Lanning, J. L. Yap, C. Gordon, P. T. Wilder, D. M. Jr, S. Fletcher, *Org. Lett.* **2013**, *15*, 3234–3237.

[230] D. S. Kemp, Z. Q. Li, *Tetrahedron Lett.* **1995**, *36*, 4175–4178.

[231] R. Larouche-Gauthier, C. J. Fletcher, I. Couto, V. K. Aggarwal, *Chem. Commun.* **2011**, *47*, 12592–12594.

[232] F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson, W. C. Still, *J. Comput. Chem.* **1990**, *11*, 440–467.

[233] A. K. Ghose, E. P. Jaeger, P. J. Kowalczyk, M. L. Peterson, A. M. Treasurywala, *J. Comput. Chem.* **1993**, *14*, 1050–1065.

[234] T. Schlick, M. Overton, *J. Comput. Chem.* **1987**, *8*, 1025–1039.

[235] J. W. Ponder, F. M. Richards, *J. Comput. Chem.* **1987**, *8*, 1016–1024.

[236] T. a Halgren, *J. Comput. Chem.* **1996**, *17*, 490–519.

[237] R. W. Hoffmann, *Angew. Chem. Int. Ed.* **1992**, *31*, 1124–1134.

[238] L. G. Ferreira, R. N. Dos Santos, G. Oliva, A. D. Andricopulo, *Molecules* **2015**, *20*, 13384–13421.

[239] D. B. Kitchen, H. Decornez, J. R. Furr, J. Bajorath, *Nat. Rev. Drug Discov.* **2004**, *3*, 935–949.

[240] S. Y. Huang, X. Zou, *Int. J. Mol. Sci.* **2010**, *11*, 3016–3034.

[241] E. Yuriev, J. Holien, P. A. Ramsland, *J. Mol. Recognit.* **2015**, *28*, 581–604.

[242] S. Forli, R. Huey, M. E. Pique, M. F. Sanner, D. S. Goodsell, A. J. Olson, *Nat. Protoc.* **2016**, *11*, 905–919.

[243] O. Trott, A. J. Olson, *J. Comput. Chem.* **2010**, *31*, 455–461.

[244] S. F. Sousa, P. A. F. M. Ramos, *PROTEINS Struct. Funct. Bioinforma.* **2006**, *65*, 15–26.

[245] M. M. Jaghoori, B. Bleijlevens, S. D. Olabarriaga, *J. Comput. Aided. Mol. Des.* **2016**, *30*, 237–249.

[246] M. W. Chang, C. Ayeni, S. Breuer, B. E. Torbett, *PLoS One* **2010**, *5*, 1–9.

[247] A. Castro-Alvarez, A. M. Costa, J. Vilarrasa, *Molecules* **2017**, *22*, 2–14.

[248] Z. Wang, H. Sun, X. Yao, D. Li, L. Xu, Y. Li, S. Tian, T. Hou, *Phys. Chem. Chem. Phys.* **2016**, *18*, 12964–12975.

[249] S. Uhrinova, D. Uhrin, H. Powers, K. Watt, D. Zheleva, P. Fischer, C. McInnes, P. N. Barlow, *J. Mol. Biol.* **2005**, *350*, 587–598.

[250] L. M. Espinoza-Fonseca, J. G. Trujillo-Ferrara, *Biopolymers* **2006**, *83*, 365–373.

[251] O. Schon, A. Friedler, S. Freund, A. R. Fersht, *J. Mol. Biol.* **2004**, *336*, 197–202.

[252] S. Yadahalli, J. Li, D. P. Lane, S. Gosavi, C. S. Verma, *Sci. Rep.* **2017**, *7*, 1–12.

[253] J. J. Sutherland, R. K. Nandigam, J. A. Erickson, M. Vieth, *J. Chem. Inf. Model.* **2007**, *47*, 2293–2302.

[254] S. Shityakov, C. Förste, *Adv. Appl. Bioinforma. Chem.* **2014**, *7*, 1–14.

[255] C. Garcia-Echeverria, P. Chene, M. J. J. Blommers, P. Furet, *J. Med. Chem.* **2000**, *43*, 3205–3208.

[256] K. Sakurai, C. Schubert, D. Kahne, *J. Am. Chem. Soc.* **2006**, *128*, 11000–11001.

[257] A. M. Beekman, M. M. D. Cominetti, S. J. Walpole, S. Prabhu, M. A. O'Connell, J. Angulo, M. Searcey, *Chem. Sci.* **2019**, *10*, 4502–4508.

[258] T. Bootwicha, J. M. Feilner, E. L. Myers, V. K. Aggarwal, *Nat. Chem.* **2017**, *9*, 896–902.

[259] Y. Gan, T. A. Spencer, *J. Org. Chem.* **2006**, *71*, 5870–5875.

[260] D. J. Pasto, R. T. Taylor, in *Org. React.*, **2004**.

[261] C. Gaviglio, F. Doctorovich, *J. Org. Chem.* **2008**, *73*, 5379–5384.

[262] Z. Lim, P. J. Duggan, S. S. Wan, G. Lessene, A. G. Meyer, K. L. Tuck, *Tetrahedron* **2016**, *72*, 1151–1160.

[263] N. M. R. McNeil, D. J. Press, D. M. Mayder, P. Garnica, L. M. Doyle, T. G. Back, *J. Org. Chem.* **2016**, *81*, 7884–7897.

[264] Y. Wang, A. Noble, E. L. Myers, V. K. Aggarwal, *Angew. Chem. Int. Ed.* **2016**, *55*, 4270–4274.

[265] H. Yin, G. I. Lee, S. P. Hyung, G. A. Payne, J. M. Rodriguez, S. M. Sebti, A. D. Hamilton, *Angew. Chem. Int. Ed.* **2005**, *44*, 2704–2707.

[266] B. B. Snider, R. A. H. F. Hui, Y. S. Kulkarni, *J. Am. Chem. Soc.* **1985**, *107*, 2194–2196.

[267] P. E. Gormisky, M. C. White, *J. Am. Chem. Soc.* **2011**, *133*, 12584–12589.

[268] K. M. Sadhu, D. S. Matteson, *Organometallics* **1985**, *4*, 1687–1689.

[269] G. Molineux, *Cancer Treat. Rev.* **2002**, *28*, 13–16.

[270] J. M. Harris, R. B. Chess, *Nat. Rev. Drug Discov.* **2003**, *2*, 214–221.

[271] M. Swierczewska, K. C. Lee, S. Lee, M. Swierczewska, K. C. Lee, S. Lee, *Expert Opin. Emerg. Drugs* **2015**, *20*, 531–536.

[272] A. K. Mailyan, K. Young, J. L. Chen, B. T. Reid, A. Zakarian, *Org. Lett.* **2016**, *18*, 5532–5535.

[273] J. Palà-Pujadas, F. Albericio, J. B. Blanco-Canosa, *Angew. Chem. Int. Ed.* **2018**, *57*, 16120–16125.

[274] G. Bifulco, P. Dambruoso, L. Gomez-Paloma, R. Riccio, *Chem. Rev.* **2007**, *107*, 3744–3779.

[275] S. Zhong, PhD Thesis. University of Bristol. Combining Computational Modelling, NMR Spectroscopy,

and Assembly-Line Synthesis for Studying Molecular Conformations., **2019**.

[276] W. Deng, J. R. Cheeseman, M. J. Frisch, *J. Chem. Theory Comput.* **2006**, *2*, 1028–1037.

[277] G. Palermo, R. Riccio, G. Bifulco, *J. Org. Chem.* **2010**, *75*, 1982–1991.

[278] C. P. Butts, C. R. Jones, J. N. Harvey, C. P. Butts, C. R. Jones, E. C. Towers, J. L. Flynn, L. Appleby, N. Barron, A. Bagno, et al., *Chem. Commun.* **2011**, *47*, 1193–1195.

[279] P. T. Robinson, T. N. Pham, D. Uhrín, *J. Magn. Reson.* **2004**, *170*, 97–103.

[280] G. Batta, K. E. Kover, *Tetrahedron* **1991**, *47*, 3535–3544.

[281] J. Wu, P. Lorenzo, S. Zhong, M. Ali, C. P. Butts, E. L. Myers, V. K. Aggarwal, *Nature* **2017**, *547*, 436–440.

[282] S. Maity, R. K. Gundampati, T. K. S. Kumar, *Nat. Prod. Commun.* **2019**, *14*, 1–17.

[283] L. Fielding, *Prog. Nucl. Magn. Reson. Spectrosc.* **2007**, *51*, 219–242.

[284] S. B. Shuker, P. J. Hajduk, R. P. Meadows, S. W. Fesik, *Science.* **1996**, *274*, 1531–1534.

[285] J. Vaynberg, J. Qin, *Trends Biotechnol.* **2006**, *24*, 22–27.

[286] A. Bax, M. Ikura, L. E. Kay, D. A. Torchia, R. Tschudin, *J. Magn. Reson.* **1990**, *86*, 304–318.

[287] L. Fielding, *Curr. Top. Med. Chem.* **2003**, *3*, 39–53.

[288] E. R. P. Zuiderweg, *Biochemistry* **2002**, *41*, 1–7.

[289] A. Medek, P. J. Hajduk, J. Mack, S. W. Fesik, *J. Am. Chem. Soc.* **2000**, *122*, 1241–1242.

[290] W. Becker, K. C. Bhattiprolu, N. Gubensäk, K. Zangger, *ChemPhysChem* **2018**, *19*, 895–906.

[291] M. P. Williamson, *Prog. Nucl. Magn. Reson. Spectrosc.* **2013**, *73*, 1–16.

[292] J. Cavanagh, A. G. Palmer, W. J. Fairbrother, M. Rance, N. J. Skelton, M. Rance, *Protein NMR Spectroscopy: Principles and Practice*, Elsevier Science & Technology, **2006**.

[293] B. Meyer, T. Peters, *Angew. Chem. Int. Ed.* **2003**, *42*, 864–890.

[294] J. J. Ziarek, F. C. Peterson, B. L. Lytle, B. F. Volkman, *Methods Enzymol.* **2011**, *493*, 241–275.

[295] A. Furukawa, T. Konuma, S. Yanaka, K. Sugase, *Prog. Nucl. Magn. Reson. Spectrosc.* **2016**, *96*, 47–57.

[296] C. H. Hardman, R. William Broadhurst, A. R. C. Raine, K. D. Grasser, J. O. Thomas, E. D. Laue, *Biochemistry* **1995**, *34*, 16596–16607.

[297] B. E. Coggins, X. Li, A. L. McClerren, O. Hindsgaul, C. R. H. Raetz, P. Zhou, *Nat. Struct. Biol.* **2003**, *10*, 645–651.

[298] R. Stoll, C. Renner, S. Hansen, S. Palme, C. Klein, A. Belling, W. Zeslawski, M. Kamionka, T. Rehm, P. Mühlhahn, et al., *Biochemistry* **2001**, *40*, 336–344.

[299] C. Riedinger, M. E. Noble, D. J. Wright, F. Mulks, I. R. Hardcastle, J. A. Endicott, J. M. Mcdonnell, *Chem. Biol. Drug Des.* **2011**, *77*, 301–308.

[300] R. Stoll, C. Renner, P. Muhlhahn, S. Hansen, R. Schumacher, F. Hesse, B. Kaluza, R. A. Engh, W. Voelter, T. A. Holak, *J. Biomol. Nmr* **2000**, *17*, 91–92.

[301] C. G. Neochoritis, M. K. Miraki, E. M. M. Abdelraheem, E. Surmiak, T. Zarganes-Tzitzikas, B. Łabuzek, T. A. Holak, A. Dömling, *Beilstein J. Org. Chem.* **2019**, *15*, 513–520.

[302] M. Bista, S. Wolf, K. Khoury, K. Kowalska, Y. Huang, E. Wrona, M. Arciniega, G. M. Popowicz, T. A. Holak, A. Dömling, *Structure* **2013**, *21*, 2143–2151.

[303] D. Muszak, B. Łabuzek, M. Z. Brela, A. Twarda-Clapa, M. Czub, B. Musielak, E. Surmiak, T. A. Holak, *J. Mol. Struct.* **2019**, *1189*, 161–174.

[304] B. Graves, T. Thompson, M. Xia, C. Janson, C. Lukacs, D. Deo, P. Di Lello, D. Fry, C. Garvie, K. Sen Huang, et al., *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 11788–11793.

[305] E. Surmiak, C. G. Neochoritis, B. Musielak, A. Twarda-Clapa, K. Kurpiewska, G. Dubin, C. Camacho, T. A. Holak, A. Dömling, *Eur. J. Med. Chem.* **2017**, *126*, 384–407.

[306] C. Riedinger, J. A. Endicott, S. J. Kemp, L. A. Smyth, A. Watson, E. Valeur, B. T. Golding, R. J. Griffin, I. R. Hardcastle, M. E. Noble, et al., *J. Am. Chem. Soc.* **2008**, *130*, 16038–16044.

[307] N. Estrada-Ortiz, C. G. Neochoritis, A. Twarda-Clapa, B. Musielak, T. A. Holak, A. Dömling, *ACS Med. Chem. Lett.* **2017**, *8*, 1025–1030.

[308] M. Krajewski, P. Ozdowy, L. D'Silva, U. Rothweiler, T. A. Holak, *Nat. Med.* **2005**, *11*, 1135–1136.

[309] R. A. Tromans, T. S. Carter, L. Chabanne, M. P. Crump, H. Li, J. V. Matlock, M. G. Orchard, A. P. Davis, *Nat. Chem.* **2019**, *11*, 52–56.

[310] P. Ríos, T. J. Mooibroek, T. S. Carter, C. Williams, M. R. Wilson, M. P. Crump, A. P. Davis, *Chem. Sci.* **2017**, *8*, 4056–4061.

[311] C. C. Lee, J. A. MacKay, J. M. J. Fréchet, F. C. Szoka, *Nat. Biotechnol.* **2005**, *23*, 1517–1526.

[312] E. R. Gillies, J. M. J. Fréchet, *Drug Discov. Today* **2005**, *10*, 35–43.

[313] S. Choudhary, L. Gupta, S. Rani, K. Dave, U. Gupta, *Front. Pharmacol.* **2017**, *8*, 1–23.

[314] J. P. Renaud, C. W. Chung, U. H. Danielson, U. Egner, M. Hennig, R. E. Hubbard, H. Nar, *Nat. Rev. Drug Discov.* **2016**, *15*, 679–698.

[315] O. Cala, F. Guillière, I. Krimm, *Anal. Bioanal. Chem.* **2014**, *406*, 943–956.

[316] A. M. Rossi, C. W. Taylor, *Nat. Protoc.* **2011**, *6*, 365–387.

[317] M. A. Williams, *Protein-Ligand Interactions: Methods and Applications*, Springer International Publishing, **2013**.

[318] P. C. Weber, F. R. Salemme, *Curr. Opin. Struct. Biol.* **2003**, *13*, 115–121.

[319] M. C. Jecklin, S. Schauer, C. E. Dumelin, R. Zenobi, *J. Mol. Recognit.* **2009**, *22*, 319–329.

[320] E. Ko, J. Liu, L. M. Perez, G. Lu, A. Schaefer, K. Burgess, *J. Am. Chem. Soc.* **2011**, *133*, 462–477.

[321] J. M. Fletcher, K. A. Horner, G. J. Bartlett, G. G. Rhys, A. J. Wilson, D. N. Woolfson, *Chem. Sci.* **2018**, *9*, 7656–7665.

[322] F. Mohamadi, N. G. J. Richards, W. C. Guida, R. Liskamp, M. Lipton, C. Caufield, G. Chang, T. Hendrickson, W. C. Still, *J. Comput. Chem.* **1990**, *11*, 440–467.

[323] N. Maestro, Version 9.2, Schrödinger, LLC, New York, **2011**.

[324] N. A. Nikolic, P. Beak, *Org. Synth.* **1997**, *74*, 23.

[325] D. C. Kapeller, F. Hammerschmidt, *J. Org. Chem.* **2009**, *74*, 2380–2388.

[326] G. E. S. M. J. Frisch, G. W. Trucks, H. B. Schlegel, V. B. M. A. Robb, J. R. Cheeseman, G. Scalmani, M. C. B. Mennucci, G. A. Petersson, H. Nakatsuji, G. Z. X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, R. F. J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, O. K. J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, J. E. P. H. Nakai, T. Vreven, J. A. Montgomery Jr., E. B. F. Ogliaro, M. Bearpark, J. J. Heyd, J. N. K. N. Kudin, V. N. Staroverov, R. Kobayashi, S. S. I. K. Raghavachari, A. Rendell, J. C. Burant, et al., *Gaussian 09, Revis. D.01, Gaussian, Inc., Wallingford CT* **2013**.